INTEGRATED CIRCUITS LABORATORY

STANFORD ELECTRONICS LABORATORIES
DEPARTMENT OF ELECTRICAL ENGINEERING
STANFORD UNIVERSITY · STANFORD, CA 94305

TR-DXG501-82   ICL-17-79

COMPUTER-AIDED DESIGN OF INTEGRATED CIRCUIT FABRICATION PROCESSES FOR VLSI DEVICES

prepared by

J.D. Plummer
R.W. Dutton
J.F. Gibbons
C.R. Helms
J.D. Meindl
W.A. Tiller
L.A. Christel
C.P. Ho
L. Mei
K.C. Saraswat
B.E. Deal
T.I. Kamins

INTEGRATED CIRCUITS LABORATORY
STANFORD UNIVERSITY
Stanford, Ca. 94305

July 1982

Final Report for Period 15 February 1981 · 31 March 1982

DTIC
ELECTE
S
NOV 10 1982
D
E

82 11 10 055

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER DELET - TR - 79 - 0257 - 3	2. GOVT ACCESSION NO. AD-A121414	3. RECIPIENT'S CATALOG NUMBER
4. TITLE *(and Subtitle)* Computer-Aided Design of Integrated Circuit Fabrication Processes for VLSI Devices (Computer-Aided Engineering of Semiconductor Integrated Circuits)		5. TYPE OF REPORT & PERIOD COVERED Final Report 15 February 1981-31 March 1982
		6. PERFORMING ORG. REPORT NUMBER 2-DXG501
7. AUTHOR(s) T. Akinwande, R. Barton, G. Bronner, L. Christel, B. Deal, W. Dibble, S. Dunham, R. Dutton, J. Gibbons, J. Greenfield, C. Han, S. Hansen, (contd)		8. CONTRACT OR GRANT NUMBER(s) MDA903-79-C-0257
9. PERFORMING ORGANIZATION NAME AND ADDRESS Stanford University Integrated Circuits Laboratory Stanford, California 94305		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Defense Advanced Research Projects Agency 1400 Wilson Blvd. Arlington, VA 22209		12. REPORT DATE July 1982
		13. NUMBER OF PAGES 444
14. MONITORING AGENCY NAME & ADDRESS(*if different from Controlling Office*)		15. SECURITY CLASS. *(of this report)* Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for public release ; distribution unlimited

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

This contract is awarded under Basic Agreement MDA903-79-C-0257, issued by Defense Supply Service-Washington. This research is sponsored by Defense Advanced Research Projects Agency (DARPA).

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

Modeling Semiconductor Devices, Semiconductor Processing, Computer-Aided Device Design, Computer-Aided Process Design, Ion Implantation, Thermal Oxidation, Chemical Vapor Deposition, Device Simulation

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

Efficient design of high performance VLSI processes requires accurate models for the physical processes used for fabrication. This is particularly true as device geometries shrink and fabrication technologies become inherently 2D. First order models for thermal oxidation, ion implantation, diffusion, chemical vapor deposition and other processes cannot accurately predict device structures which result from modern IC technologies. (Continued over)

DD FORM 1473 EDITION OF 1 NOV 65 IS OBSOLETE
1 JAN 73

20.   (Continued)

The fundamental objective of this program is to develop accurate and physically correct models for these processes which are general enough to incorporate in a general purpose, user-oriented computer simulation tool - SUPREM. This program accepts process schedules as inputs and provides predicted device structures as outputs. It is meant to be capable of accurately simulating both bipolar and MOS VLSI structures. SUPREM is specifically designed to couple with device simulation tools so that it forms the cornerstone of a hierarchy of VLSI process, device, circuit and system design aids.

7. Cont'd--
C. Helms, C. Ho, T. Kamins, M. Kump, E. Kutlu, L. Lie, H. Massoud, L. Mei, J. Meindl, D. Modlin, J. Plummer, R. Razouk, J. Rouse, K. Saraswat, F. Shone, B. Swaminathan, Y. Thathachari, W. Tiller, M. Taubenblatt, E. Young.

FINAL REPORT

ON

COMPUTER-AIDED DESIGN OF INTEGRATED CIRCUIT
FABRICATION PROCESSES FOR VLSI DEVICES

15 FEBRUARY 1981 - 31 MARCH 1982

This work was supported by

a contract from the

Accession For	
NTIS GRA&I	X
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A	

James D. Plummer

Stanford Electronics Laboratories

Stanford University   Stanford, California

TABLE OF CONTENTS

COMPUTER AIDED DESIGN OF INTEGRATED CIRCUIT

FABRICATION PROCESSES FOR VLSI DEVICES

Introduction

The work described here builds upon an established program which over the past five years has resulted in the development of three generations of computer aided process modeling programs - SUPREM I, II and III. The second generation version - SUPREM II - is currently in use at more than 150 industrial firms and university research groups and is generally regarded as the pre-eminent tool of its type available today. A significant number of published papers from groups all over the world reference this computer program, and in fact, many papers include simulated results obtained from SUPREM. It has found widespread application in the design of state-of-the-art MOS and bipolar fabrication processes.

SUPREM III, which will be released in the immediate future, substantially upgrades the program in that it makes it possible to simulate structures with up to five material layers. Built in models for Si, SiO_2, Polysilion and Si_3N_4 are included and other materials can easily be added with user defined parameters. A large number of new and/or improved physical models for oxidation, diffusion, epitaxy and ion implantation are included in SUPREM III, these have, for the most part, come from experimental and theoretical work under this program over the past five years. These new models should greatly expand the range of applicability and the accuracy of the program.

The specific goals of this research have been and continue to be the formulation of basic physical and mathematical models of integrated circuit

1

fabrication processes which can accurately predict the structure of small geometry VLSI devices (and other arbitrary device structures) which will result from a given fabrication sequence, and to implement these models in a comprehensive process modeling computer program. Such a program, it was believed at the outset of this work, would significantly reduce costly and time consuming iterative, experimental approaches to developing or optimizing silicon technologies. This has in fact occurred. SUPREM III represents a major step along the path towards improved simulation tools.

It is clear that the historical trends in integrated circuit technology over the past two decades towards increased complexity and smaller active device dimensions will continue. Modern device structures employ lateral device dimensions on the order of 2-3 μ and vertical dimensions well below 1 μm. There are no basic physical mechanisms which will prevent a reduction in each of these dimensions by an additional order of magnitude over the next decade. Practically accomplishing this continued scaling, however, depends upon our ability to physically understand and quantitatively model the fabrication techniques which will be used in the construction of such devices.

For devices with relatively large geometries (>5 μ) and loose processing tolerances, relatively simple models suffice for prediction of vertical device structures resulting from a given fabrication sequence. As device dimensions shrink, however, it becomes essential to employ more robust process models and to consider the interaction both laterally and vertically of various processing steps, if accurate simulation of structures is to be obtained. This is important even with today's 2-3 μm device geometries; it will become essential for smaller devices.

Large geometry devices can be successfully modeled as one-dimensional structures. This is true for both process models and electrical models.

Devices with lateral dimensions below a few microns, however, require two-dimensional models for accurate simulation. This need has stimulated a large body of work in recent years on two-dimensional electrical models of device current-voltage characteristics. One example of this type of work is the 2D program GEMINI, developed under this program. Work of this type has resulted in remarkable advances in our understanding of small geometry device physics.

Progress has not been as rapid, however, in two-dimensional process modeling. While some basic 2D process modeling programs have been developed at Stanford (SUPRA for example) and elsewhere, these programs do not incorporate robust 2D kinetic models which can accurately predict doping profiles and device geometries under a wide range of conditions. This is a direct result of our need for improved physical models of oxidation, ion implantation, diffusion, and CVD. It is quite clear that these processes are not one dimensional. Recent experimental evidence has clearly indicated that oxidation or impurity diffusion in a localized region of a silicon substrate can substantially affect oxidation or diffusion rates in laterally or vertically adjacent regions of the substrate. There is no clear agreement at the present time on the basic physical mechanisms responsible for such results, although much has been learned in the past several years. It is clear, however, that we must quantitatively understand such phenomena if we are to accurately model small device structures.

A specific goal of this program is to understand and model these two-dimensional effects. We have made substantial progress in this regard in the past year. It appears now that the basic physical phenomena underlying these interactions are the roles of point defects--silicon vacancies and interstitials--in impurity diffusion, thermal oxidation, and other processes. The generation and consumption of these point defects during high temperature

3

fabrication steps appear to be the unifying physical effects which can explain many of the phenomena which have been regarded as anomalous to date. We have used such models to quantitatively understand a variety of process phenomena and have incorporated some of these models in SUPREM II. More such models are in SUPREM III. The unifying role of these mechanisms is a cornerstone of the work described in this report. We regard such models as absolutely essential to accurate modeling of two-d.mensional effects in small devices.

Simply stated, the long-range objective of this research is the development of a 2D process modeling program which contains physically correct 2D kinetic models. Most of the effort under this program will be aimed at experimental and theoretical work to uncover the basic physical mechanisms which govern 2D oxidation kinetics, 2D diffusion, etc. This is a major problem area with very difficult materials problems. It will become clear in the technical discussions later in this report that 2D process modeling involves much more than solving 1D equations in two dimensions. Basic physical laws and understanding are lacking at present. Effects such as lateral OED, oxidation under masking Si_3N_4 layers, lateral and vertical effects of high dopant concentrations on diffusion coefficients, and a host of other known experimental effects cannot be explained by a simple extrapolation of known 1D physical laws to 2D structures. This program is aimed at generating the physical understanding needed to develop 2D kinetic models.

An essential part of this will be to develop models for bulk point defect (interstitial and vacancy) generation, recombination and diffusion, since it is clear that the local concentrations of these defects determine local diffusion coefficients, oxidation rates, etc. In fact an alternative statement of the overall objective of this program would be to develop techniques for calculating local (i.e. time and position dependent) process parameters

4

suitable for process simulation. Such process parameters will of necessity be
geometry dependent which means that diffusion coefficients and oxidation rate
constants, for example, will depend on the presence or absence of nearby heav-
ily doped regions or oxidizing interfaces. This will imply a tight coupling
between surface geometry and resulting impurity profiles in small devices.

In this context, the overall objectives of this program are:

(1) To develop accurate physical models for the basic technologies
used in the fabrication of silicon integrated circuits--oxidation,
ion implantation, diffusion, and chemical vapor deposition (CVD).

(2) To develop accurate physical models for the interaction of these
processes, for example, the enhancement of diffusion coefficients
in an oxidizing ambient.

(3) To develop accurate physical models based upon silicon point defects
for two-dimensional process phenomena in small geometry devices.

(4) To develop high resolution analytic tools for the experimental
characterization of small device structures.

(5) To implement improved process models in updated versions of SUPREM III
so that this program is capable of accurately simulating small devices
and multi-layer structures.

(6) To aim toward a robust 2D process modeling program (SUPREM IV) which
would combine outputs from this program with numerical techniques
developed under other support.

(7) To disseminate the results to the defense industry.

Because this is a fairly mature research program, the next section
describes our present situation with regard to SUPREM and its process models
in summary form. This section is intended to provide an overall perspective

of the program and to provide background information for the detailed sections which follow. These later sections describe specific activity in the past year and project proposed activity in the future.

1. PRESENT STATUS OF SUPREM

The overall objective of the process modeling program SUPREM is to permit a process/device designer to accurately simulate complete silicon fabrication technologies. The program input is, in essence, a processing schedule specifying a sequence of times, temperatures, ambients, depositions, implant doses and energies, predepositions, etc. The program output, available after each step in the process sequence, is, in the case of SUPREM I and II, the one dimensional impurity profile in the bulk silicon and any overlying SiO_2 layer. In the case of SUPREM III, the structure may incorporate additional material layers above the bulk and/or SiO_2 including polysilicon and Si_3N_4. The thicknesses and dopant profiles in each of these layers are calculated by the program during each process step.

By way of introduction to the work described in this report, it will be useful to review the past achievements and present status of the physical models used in SUPREM. We will do this by illustrating the progression of physical models from SUPREM I through SUPREM III. The increase in capability of the program through each of its versions should become apparent.

The basic structure of SUPREM is illustrated in Fig. 1.1. The program is designed so that steps can be simulated either individually or sequentially just as they would occur during the actual fabrication of an IC. In addition to dopant profiles and layer thicknesses, junction depths, sheet resistance and MOS threshold voltages are also calculated in versions II and III. An example of a typical SUPREM output in this case from an early version of SUPREM III is shown in Fig. 1.2. In the following pages, a detailed discussion of the models used in versions I, II and III will be given.

1.1 THERMAL OXIDATION

The rate of SiO_2 growth on silicon has historically been described by the well-known linear-parabolic growth law [1.1].

$$\frac{X_o^2}{B} + \frac{X_o}{B/A} = t + \tau \qquad (1.1)$$

where X_o is the oxide thickness, t is time, and B and B/A are the parabolic and linear rate constants respectively. B is proportional to the oxidant solubility in the SiO_2 (and hence to oxidant partial pressure), and to the oxidant diffusion coefficient in the SiO_2. B/A is also, to first order, proportional to oxidant solubility (and hence to pressure) and to k_s, the Si/SiO_2 interface reaction rate constant. The two rate constants B and B/A separate the overall oxidation process into diffusion of the oxidant species through any existing SiO_2 layer and reaction at the Si/SiO_2 interface, either one of which may be rate limiting depending on temperature and SiO_2 thickness.

The parameter τ is related to any initial oxide thickness, X_i at $t = 0$ by

$$\tau = \frac{X_i^2 + AX_i}{B} \qquad (1.2)$$

and has also been used to correct eqn. 1.1 for the "anomalous" fast inital growth rate up to $X_o = 200$ Å in dry O_2.

As written in equations 1.1 and 1.2, the oxidation growth law provides limited physical insight into the underlying physical processes involved. One can generally state that anything which affects the rate of diffusion of O_2 or H_2O through the SiO_2 layer (such as high dopant concentrations in the SiO_2) should affect B. Alternatively, anything which affects the Si/SiO_2 interface chemical reaction rate (such as substrate crystal orientation) should affect B/A. However, the detailed mechanisms which may be responsible for these changes in B and B/A are not apparent from eqn. 1.1.

In modern device structures, which are making increasing use of lower temperature processes and thinner SiO_2 layers, B/A or the interface reaction

8

rate is increasingly the dominant rate limiting parameter. It has also become clear in recent years that the detailed mechanisms involved in B/A are responsible for many important process phenomena (oxidation enhanced diffusion or OED for example). A substantial amount of our work in thermal oxidation over the past several years has, therefore, concentrated on determining the detailed mechanisms involved in B/A. This activity has been reported in detail in our previous annual reports. Since it is essential to much of the work to be described later, we outline here the basic ideas which have evolved from this work.

The basis of our model is shown in Fig. 1.3. O_2 from the gas phase is incorporated into the surface of the SiO_2 layer and is known to diffuse to the Si/SiO_2 interface where the reaction converting silicon to SiO_2 takes place. The work of Doremus [1.2], Jorgensen [1.3], Rayleigh [1.4], and Collins and Nakayama [1.5] presented conflicting evidence concerning the presence or absence of an important charged oxygen ion in the diffusion process. The work of Mills and Kroger [1.6] clearly showed that a doubly negative oxygen interstitial ion was the dominant species in the high temperature conductivity of SiO_2. The first-order pressure dependence of the parabolic rate constant B in the Deal and Grove model [1.1] seems to indicate that the dominant diffusing species is molecular oxygen. It is quite surprising that 20 years of work on this basic question has not resolved the issue of the main diffusing species in Si oxidation. For the purposes of the present discussion, we will consider O_2 as the dominant species for thick oxide layers.

At the Si/SiO_2 interface, dissociation of the O_2 occurs as a first step in the oxidation reaction. When silicon and an oxygen species react in the interface region to form an SiO_2 molecule, a necessary quantity of "free volume" must also be supplied so that the molecule can fit into the normal

9

SiO_2 network structure. From density considerations, the average spacing between Si atoms in SiO_2 is about 1.3 times the average spacing between Si atoms in the Si lattice. This would lead to a 70% strain in the SiO_2 if the lattice continued in a coherent way into the SiO_2 film. If the interface moves into the Si at V cm/sec, the oxide thickens at 2.25 V cm/sec, which means that the "free volume" must flow in at a rate of 1.25 V cm/sec per cm^2 of interface area, to produce an unstrained film. If no "free volume" is supplied, then the excess free energy stored as strain in the SiO_2 film is sufficient to strongly retard the oxidation reaction at normal driving forces.

A portion of the required "free volume" for the reaction may be supplied as shown in the top reaction in Fig. 1.3. In general, we might expect some lattice mismatch to occur at the interface, stored as a cross-grid of dislocation lines (dangling bonds) and some of the mismatch to be stored as strain in the SiO_2 film.

Two alternative means of providing the necessary free volume at the interface are shown in the middle and bottom of Fig. 1.3 and are based upon point defects in the SiO_2 and silicon layers. In the middle reaction, vacancies (Si_V) from the silicon substrate provide the reaction sites or the required "free volume." In the bottom reaction, silicon atoms are removed from lattice sites to create interstitials (Si_I) and the required reaction sites for the growth of the SiO_2. Note that possible charge states of the point defects have been neglected in the reactions proposed here. Si vacancies are known to exist in at least three charge states, Si_V^+, Si_V^-, and $Si_V^=$, in addition to neutral vacancies, very little is known at the present time about possible Si_I charge states.

It is believed that all three of these mechanisms play a role in the oxidation process. Each of them likely dominates under specific process conditions. For example, the reaction involving Si_V is believed to dominate in the case of

heavily doped N^+ substrates when the position of the Fermi level is near the conduction band and the equilibrium concentration of vacancies is dramatically increased. This mechanism is the basis for our quantitative model explaining the enhanced oxidation rates of N^+ substrates [1.7,1.8].

The bottom mechanism in Fig. 1.3 has been postulated by Dobson [1.9,1.10] and the excess interstitial flows produced during thermal oxidtion may be the mechanism of oxidation enhanced dopant diffusion effects [1.9-1.12]. In addition, the interstitials produced during the reaction have been linked to oxidation-induced stacking faults [1.11, 1.13]. Finally, it might be conjectured at this point that some of the excess interstitials move into the SiO_2 layer. Such a flow would then very probably be closely related to the observed fixed oxide charge N_f that has been discussed as incompletely oxidized silicon atoms [1.14].

This basic concept of the roles of point defects in many related processes has been successfully used to quantitatively model the effects of enhanced oxidation over heavily doped regions [1.7,1.8]; to quantitativley model the growth and retrograde growth of oxidation induced stacking faults [1.11,1.13]; to qualitatively explain enhanced diffusion coefficients during thermal oxidation [1.12]; and to qualitatively explain the dependence of fixed oxide charge (N_f) on process parameters [1.15]. These and other implications of this type of modeling are indicated in Fig. 1.4. Such a model thus appears to be of major significance in accurately modeling and understanding many of the fabrication steps used in silicon integration circuits. It is believed that as device geometries shrink towards submicron structures, this type of point defect modeling will be essential if we are to accurately model VLSI structures. Additional discussion of this model and its implications are contained in our last two annual reports. [1.15,1.16].

11

We turn now to a detailed discussion of the oxidation kinetic models used in SUPREM I, II and III. These models are summarized in Table I which will be referred to in the following paragraphs. All three versions of SUPREM produce 1D vertical cross-sections through the illustrated structures in the Table.

All versions of SUPREM employ an incremental form of the classic linear-parabolic growth equation. The increment of oxide thickness ΔX_0 grown from time point t_{i-1} to t_i is calculated as

$$\Delta X_0 = X_i - X_{i-1}$$

$$= \frac{1}{2} \left[-(2X_{i-1} + A) + \sqrt{(2X_{i-1} + A)^2 + 4B(t_i - t_{i-1})} \right] \quad (1.3)$$

For each successive increment, the rate constants B and B/A apropriate for the substrate and ambient conditions applicable to that interval are calculated. Specifically incorporated in the functional format of the rate constants are numerous process variables demonstrated to influence oxidation kinetics that have been quantitatively characterized and modeled under this program. Thus, for either H_2O or dry O_2 oxidations,

$$B = B^i \left[\frac{1}{2} (P_i + P_{i-1})\right] \left[1 + \delta C_T^q\right] \left[\epsilon\right] \quad (1.4)$$

— HCL dependence

— substrate doping dependence

— oxidant pressure dependence

— "intrinsic" parabolic rate constant
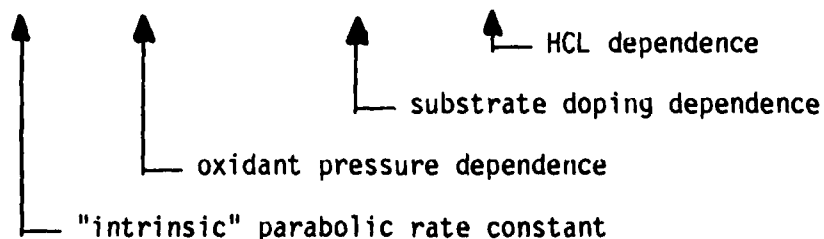
Similarly, the linear rate constant for H_2O oxidation is

12

$$B/A = (B/A)^i \left[\tfrac{1}{2}(P_i + P_{i-1})\right] \left[1 + \vartheta(C^T - 1)\right] [\eta] [\alpha] \qquad (1.5)$$

— orientation dependence

— HCl dependence

— substrate doping dependence

— oxidant pressure dependence

— "intrinsic" linear rate constant

However, for dry O_2 oxidations, B/A has been found to have a different oxidant pressure dependence, and an "anomalous" fast initial oxidation as discussed below; thus,

$$B/A = (B/A)^i \left[\tfrac{1}{2}(p_i^{0.75} + p_{i-1}^{0.75})\right] \left[1 + \vartheta(C^T - 1)\right] [\eta] [\alpha] \left[1 + Ke^{-X_i/\ell}\right] \quad (1.6)$$

— thin oxides

Equations 1.4-1.6 express the most general relationship implemented in SUPREM (Version III) to describe oxidation kinetics. Versions I and II contain only a subset of the factors in eqns. 1.4-1.6 as described below.

1.1.A.   Lightly Doped "Intrinsic" Kinetics

Rate constants for the oxidation of lightly doped, <111>-oriented single crystal silicon at one atmosphere oxidant pressure (no chlorine additions to the ambient) were determined in the first year or two of this program. For dry O_2, both B^i and $(B/A)^i$ may be well-represented as singly-activated processes [1.17,1.18]:

$$B^i = C_1 e^{-E_1/kT} \qquad \text{where} \qquad C_1 = 7.72 \times 10^2 \ \mu^2/hr \qquad (1.7)$$
$$E_1 = 1.23 \ eV$$

$$(B/A)^i = C_2\, e^{-E_2/kT} \quad \text{where} \quad C_2 = 6.23 \times 10^6 \ \mu/hr \tag{1.8}$$

$$E_2 = 2.0 \ eV$$

For H_2O oxidation, both B^i and $(B/A)^i$ also appear to be activated processes but with differing activation energies for higher vs. lower temperature ranges [1.15, 1.19]:

$$B^i = C_1\, e^{-E_1/kT} \quad \text{where } T > 950°C: \qquad C_1 = 4.20 \times 10^2 \ \mu^2/hr \tag{1.9}$$

$$E_1 = 0.78 \ eV$$

$$T < 950°C: \qquad C_1 = 1.70 \times 10^4 \ \mu^2/hr$$

$$E_1 = 1.17 \ eV$$

$$\left(\frac{B}{A}\right)^i = C_2 e^{-E_2/kT} \quad \text{where } T > 900°C: \qquad C_2 = 1.77 \times 10^8 \ \mu/hr \tag{1.10}$$

$$E_2 = 2.05 \ eV$$

$$T < 900°C: \qquad C_2 = 2.07 \times 10^6 \ \mu/hr$$

$$E_2 = 1.60 \ eV$$

"Pyrogenic steam" and "wet O_2" (a 95°C H_2O bubbler) in one atmosphere ambient pressure have been estimated to correspond to an effective H_2O oxidant pressure $P_{H_2O} = 0.92$ atm [1.19]. The rate constants will vary accordingly, as discussed below.
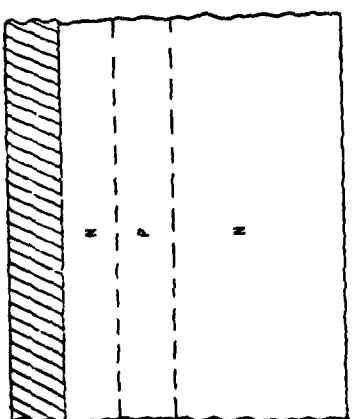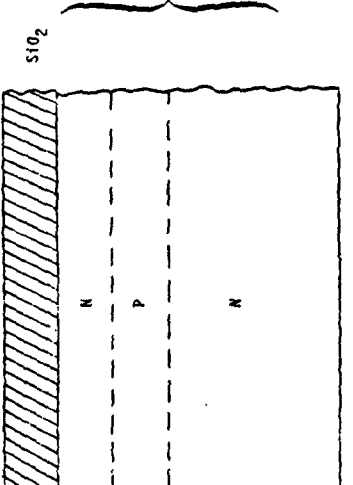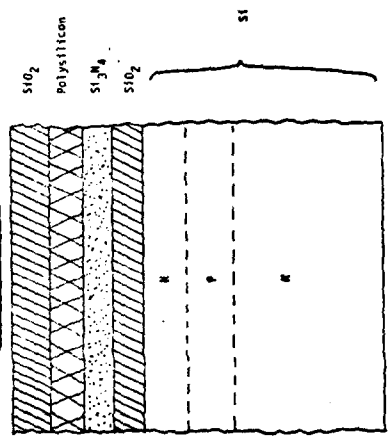
B has been found to have no significant substrate orientation dependence. However, B/A is observed to be ~ 1.7X greater for <111> than for <100> silicon in both H_2O and dry O_2 oxidations [18]. Thus,

$$\propto \langle 111 \rangle = 1 \tag{1.11a}$$

$$\propto \langle 100 \rangle = 1/1.7 = 0.595 \tag{1.11b}$$

Dependencies for other orientations have not been determined, although they are expected to lie between the <111> and <100> rates.

14

Table 1.1 - THERMAL OXIDATION MODELS

SUPREM 1 — SiO_2 on Si (N, P, N)
- 2 layers
- 1 interface
- 2 materials

SUPREM II — SiO_2 on Si (N, P, N)
- 2 layers
- 1 interface
- 2 materials

SUPREM III — SiO_2, Polysilicon, Si_3N_4, SiO_2, Si (N, P, N)
- 5 layers
- 4 interfaces
- Any materials

All Versions:

$$\frac{(X - X_i)^2}{B} + \frac{(X - X_i)}{B/A} = t$$

SUPREM I

$$\frac{B}{A} = \frac{B}{A}^i \{(P)\} \; [1 + \gamma(C^T - 1)] \; (\alpha)$$

Pressure   Doping   Orientation

$$B = (B)^i \; (P) \; (1 + \delta C_T^q)$$

Pressure Doping

$$\frac{B}{A}^i = C_1 e^{-E_1/kT}$$

$$(B)^i = C_2 e^{-E_2/kT}$$

SUPREM II

$$\frac{B}{A} = \frac{B}{A}^i \{(P) \; [1 + \gamma(C^T - 1)] \; (\alpha)\}$$

Pressure   Doping   Orientation

$$B = (B)^i \; (P) \; (1 + \delta C_T^q)$$

Pressure Doping

SUPREM III

$$\frac{B}{A} = \frac{B}{A}^i \{(P)^n [1 + \gamma(C^T - 1)] \; (n) \; (\alpha) \; (1 + Ke^{-\frac{X}{L}})\}$$

Pressure   Doping   HCl   Orientation   Thin Oxides Dry O_2

$$B = (B)^i \{(P)^m \; (1 + \delta C_T^q) \; (\epsilon)\}$$

Pressure Doping   HCl

15

As indicated in Table 1.1, SUPREM I's complete oxidation model consisted of eqns. 1.7-1.11. In fact the break in activation energy reflected in equations 1.9 and 1.10 in B and B/A for H_2O oxidation, was not included in SUPREM 1. Versions II and III contained all of the above models for "intrinsic" oxidation.

1.1.B. Oxidant Pressure Dependence

Recent work under this program has characterized the effect of oxidant pressure on the kinetics of silicon oxidation in both dry O_2 and pyrogenic steam for $T = 800-1000°C$ and $P = 1-20$ atm. In pyrogenic steam, both linear and parabolic rate constants were found to be linearly proportional to H_2O pressure as indicated in Fig. 1.5 [1.15,1.19]. The same relationship was observed by Deal and Grove for pressures below 1 atm [1.1].

For dry O_2, B also exhibited a linear pressure dependence, as in Fig. 1.6. However, the linear rate constant B/A demonstrated a sub-linear pressure dependence, with proportionality to the power of n where $0.7 < n < 0.8$, as shown in Fig. 1.7 [1,16, 1.20].

Typical high pressure oxidation cycles employ an approximately linear pressure ramp in the oxidizing ambient at temperature to acheive the desired oxidation pressure. Addition of a linear pressure ramp capability in SUPREM allows simulation of the complete cycle oxidation kinetics. For such a linear pressure ramp, the effective rate constants during the time increment t_{i-1} to t_i may be calculated. In the case of a linear pressure dependence of the rate constants, as for both B and B/A in H_2O and B in dry O_2 oxidations, the effective rate constant is found to be simply the average value during that time increment:

$$B = B_i \left[1/2(P_i + P_{i-1}) \right] \qquad (1.12)$$

In the more general case of a P^n dependence, as for B/A in dry O_2 oxidations, however, the effective rate constant becomes

$$\frac{B}{A} = (B/A)^i \left[\frac{1}{(n+1)} \quad \frac{P_i^{n+1} - P_{i-1}^{n+1}}{(P_i - P_{i-1})} \right]$$ (1.13)

This dependence may be seen to reduce to the simple average when $n = 1$. Also, when pressure is unchanged during the time interval, the dependence reduces to the proper P^n.

For ease of implementation, a simple averaging during each time increment has been incorporated in SUPREM III. This approach is rigorously correct for both B and B/A in H_2O oxidation, and for B in dry O_2 oxidation. For B/A in dry O_2 kinetics, some error is introduced. This error should be small and in any case is incurred only during the pressure ramp, a small contribution to the total oxide growth. SUPREM versions I and II assumed linear pressure dependence for both B and B/A under all conditions and thus did not encounter this problem.

1.1.C. Substrate Doping Dependence

Under conditions of high substrate doping levels, oxidation rates may be increased substantially [1.21]. Work under this program [1.22, 1.7, 1.8] has suggested that the major enhancement is seen in B/A or the interface reaction rate. Increasing substrate doping increases the total silicon substrate vacancy concentrations [1.23]. Vacancy consumption has been suggested to be directly involved in the interface oxidation reaction as described earlier. Thus the linear rate constant has been shown to correlate directly with the increased vacancy levels with heavy doping,

$$B/A = (B/A)^i \left[1 + \mathit{v}(C^V - 1) \right]$$ (1.14)

17

where ୪ is a measure of the intrinsic vacancy role in the interface oxidation reaction, experimentally found to be

$$\text{୪} = 2.62 \times 10^3 \ \exp\left(\frac{-1.10 \ eV}{kT}\right) \tag{1.15}$$

and C_V is the normalized total vacancy concentration, given by

$$C_V = \frac{1 + C^+\left(\dfrac{n_i}{n}\right) + C^-\left(\dfrac{n}{n_i}\right) + C^=\left(\dfrac{n}{n_i}\right)^2}{1 + C^+ + C^- + C^=} \tag{1.16}$$

with

$C^+ = \exp\left[(E^+ - E_i)/kT\right],$ $\qquad\qquad E^+ = 0.35 \ eV$

$C^- = \exp\left[(E_i - E^-)/kT\right],$ $\qquad\qquad E^- = E_g - 0.57 \ eV$

$C^= = \exp\left[2E_i - E^- - E^=)/kT\right],$ $\qquad\quad E^= = E_g - 0.12 \ eV$

The above expressions should be recognized as the normalized intrinsic concentrations of vacancies in the three charge states with their corresponding state energies in the silicon bandgap. Finally, the silicon energy bandgap Eg, intrinsic Fermi level E_i, and intrinsic carrier concentration n_i are

$$E_g = 1.17 - 4.73 \times 10^{-4} \left[\frac{T^2}{T + 636}\right] \ eV \tag{1.17}$$

$$E_i = E_g/2 + \frac{3kT}{4} \ (\ell n \ 0.719) \tag{1.18}$$

$$n_i = 3.87 \times 10^{16} \ T^{3/2} \ \exp\left(\frac{-0.605 \ eV}{kT}\right) \ cm^{-3} \tag{1.19}$$

For phosphorus doping, a small enhancement of the parabolic oxidation rate has also been measured. This is likely due to an increased oxidant diffusivity in the SiO$_2$ due to dopant segregation into the SiO$_2$. B is modeled empirically as

18

$$B = B^i \left[1 + \delta C_T^q \right] \qquad (1.20)$$

where

$$\delta = 9.63 \times 10^{-18} \quad \exp\left(\frac{+2.83 \text{ eV}}{kT}\right) \qquad (1.21)$$

$$q = 1.28 \exp\left(\frac{-0.176 \text{ eV}}{kT}\right) \qquad (1.22)$$

C_T is the total concentration of dopant in the substrate side of the Si/SiO$_2$ interface (but not the interface pile-up peak level seen with phosphorus). A "threshold" doping level of ~10^{19} cm^{-3} should perhaps be reached before eqn. (1.21) is applied.

Eqn. (1.21) should also be reasonably applicable for other n-type dopants such as arsenic and antimony that should show fairly similar behavior in segregation to and complexing in the growing oxide. For p-type dopants such as boron, however, the dependence of B on doping level is even less well char·c-terized. Irene and Dong [1.24] have found for oxidation in dry O$_2$ of a single heavy boron doping level of ~6 x 10^{19} cm^{-3} that enhancement of B is significant but with a decidedly different temperature behavior than for comparable phosphorus doping levels. Substantially more work is obviously needed for both p and n-type dopants.

Note that n-type dopants show substantial interface pile-up during oxidation (as described in previous annual reports and in Section 16 of this report). However, there is still some question as to what degree these pile-up peaks are either mobile or electrically active. In any event, the parameter values for eqns (1.15) and (1.21) were derived without considering the pile-up and peak levels but rather the doping levels just beyond the very narrow pile-up peaks (~100 Å into the silicon).

As shown in Table 1.1, SUPREM II and III contain the heavy doping model described above. SUPREM I did not.

19

1.1.D. Chlorine Additions To The Ambient (HCl Dependence)

Addition of a chlorine species to the oxidizing ambient has become common industrial practice for the passivation of thermally grown silicon dioxide films. Improved threshold stability and more uniform dielectric strength of such films have been well documented [1.25-1.27]. Significant increases of oxidation rate by such chlorine additions have also been observed [1.25, 1.28, 1.29].

Various chlorine species have been used, including most commonly HCl [1.30], trichloroethylene C_2HCl_3 (TCE) [1.31], and 1.1.1. - trichloroethane H_3CCl_3 (TCA, C33) [1.32]. Understanding and physical modeling of the kinetics for any of these species, however, have not been achieved. Perhaps the most extensive data base exists for HCl oxidation, including substantial work done under this program as described in previous years' reports [1.33,1.34] and journal papers [1.30,1.35,1.36]. While this data has not been conclusive in understanding the physical mechanisms involved, the data is of immediate value as the range of temperature and % HCL ambient includes most conditions of practical interest. This data is also directly applicable to TCA oxidations because of the identical H/Cl ratios in the two species. A given % TCA ambient has been shown to be equivalent to an HCl ambient of exactly three times that % [1.32]. No such simple direct equivalence has been demonstrated for TCE oxidation. (However, the use of TCE as the chlorine species has decreased.)

A basic modeling capability of HCl and (TCA) oxidation kinetics is therefore feasible for inclusion in SUPREM based empirically on the available HCl data. This capability, suitable for most chlorine oxidations of practical interest to the process engineer, is implemented in SUPREM III as multiplicative factors in the calculation of both B/A and B,

20

$$B/A = (B/A)^i \, \eta \qquad\qquad (1.23)$$

$$B = B^i \, \varepsilon \qquad\qquad (1.24)$$

In this formulation, η and ε contain the HCl dependencies of B/A and B, respectively, and will vary with θ = % HCl, temperature T, and major component of the ambient (O_2 vs. H_2O).

For HCl/O_2 ambients, the extensive data produced in this program includes 0 - 10% HCl for T = 900 - 1100°C. As the HCl concentration increases from 0 to 10% for a given temperature, the oxidation rate increases. This increase is reflected in a continuous increase of the parabolic rate constant B, but the linear rate constant B/A levels off with increasing HCl after an initial increase with the first 1-2%, as seen in Figs. 1.8 and 1.9 [1.30]. The increases in B/A are similar for both <100> and <111> orientations. Thus, no significant additional orientation dependence appears to be introduced by HCl, and the intrinsic orientation dependence already included in α of eqn. (1.11) remains adequate.

Because understanding the mechanisms involved is still incomplete, this data is implemented in SUPREM III as look-up tables of η and ε, as in Tables 1.2 and 1.3. Linear interpolation is employed for conditions of % HCl and T intermediate to available data points in the range of the look-up tables, for want of any strong contraindications. For conditions beyond the range of available data, the best procedure is even less obvious. Some limited data for T = 1150°C from 0 to 11% HCl has been analyzed for an effective B only [1.37]. The observed enhancement ε values are very similar to the corresponding values for T = 1000°C in Table 1.3. So in SUPREM III the last applicable table point (or interpolated value) is employed for both ε and η for conditions beyond the range of Tables 1.2 and 1.3.

21

Table 1.2

T =	900°C	1000°C	1100°C
θ = 0% HCl	η = 1.0	1.0	1.0
1%	1.750	1.250	1.621
3%	1.750	1.486	2.207
5%	1.750	1.486	2.207
7%	1.750	1.486	2.207
10%	1.750	1.486	2.207

Look-up Table for SUPREM Implementation of HCl Dependence η of Linear Rate Constant B/A on Temperature T and % HCl θ in HCl/O_2 Oxidations.

Table 1.3

T =	900°C	1000°C	1100°C
θ = 0% HCl	ε = 1.0	1.0	1.0
1%	1.083	1.658	1.355
3%	1.250	1.840	1.490
5%	1.444	2.075	1.641
7%	1.639	2.332	1.816
10%	2.028	2.759	2.102

Look-up Table for SUPREM Implementation of HCl Dependence ε of Parabolic Rate Constant B on Temperature T and % HCl θ in HCl/O_2 Oxidations.

For HCl/H2O oxidations, data produced in this contract included additions of 5% HCl for T = 900 - 1100°C. The oxidation rate is found to decrease with HCl such that addition of HCl serves to dilute the oxidizing species. The effective parital pressure of the oxidant is reduced. This reduction is calculated in SUPREM III for a given θ = % HCl via

$$\eta = 1 - (\theta/100) \qquad (1.25)$$

$$\epsilon = 1 - (\theta/100) \qquad (1.26)$$

This empirical approach should be useful for most chlorine oxidations of practical interest, within the range of conditions for which data is currently available. Better understanding of the mechanisms involved will be necessary, however, for complete predictive capability particularly outside the available data range.

The modeling capability expressed in equations 1.23-1.26 and Tables 1.2and 1.3 is implemented only in SUPREM III.

1.1.E. Thin Oxide Dry O_2 Kinetics

It is well known that the growth of very thin (<500Å) SiO_2 layers in dry O_2 occurs by fundamentally different mechanisms than thicker layers [1.1,1.16]. Direct manifestations of this include faster growth kinetics and different optical and physical properties. No satisfactory explanation has yet been provided for these results. We believe that SUPREM must be able to accurately model such thin SiO_2 layers since gate oxides <400 Å are already an essential part of VLSI MOS circuits. Because of this a major part of our oxidation activities in the past two years has been aimed at resolving these questions. We have developed special experimental techniques (in-situ etching prior to oxidation) and have gathered a substantial amount of experimental kinetic data. This body of data has been considerably expanded during the present contract year under a joint program with E. Irene of the IBM Watson Research Laboratory [1.16].

23

Some experimental data also exists in the literature on the growth kinetics of thin SiO_2 layers [1.38-1.45]. One of the first detailed studies was reported by Van der Meulen and Ghez [1.38,1.39]. Their work involved oxidations between 700° and 1000°C in dry O_2 and in O_2/N_2 mixtures with the O_2 partial pressure as low as 0.01 atm. They found that both <111> and <100> orientations exhibited essentially linear growth at least down to ~50 A in thickness. The pressure dependence of the growth rate, however, varied from close to $p^{0.5}$ at 700°C up to close to $p^{1.0}$ at temperatures of 1000°C and higher. This result is in contradiction to the predictions of the linear-parabolic growth law [1.1] which predicts a $p^{1.0}$ pressure dependence under all conditions, but agrees with the high pressure kinetic data obtained under this program and described earlier.

In an effort to explain their experimental results, Van der Meulen and Ghez proposed that the interface reaction is actually more complex than that proposed by Deal and Grove. They suggested that three reactions occur, given by

$$Si-Si + O_2 \rightarrow Si-O-Si + O \qquad (1.27)$$

$$Si-Si + O \rightarrow Si-O-Si \qquad (1.28)$$

$$O_2 \rightleftharpoons 2O \qquad (1.29)$$
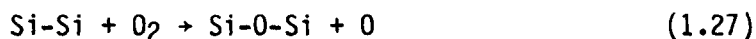
While the first two reactions are not stoichiometric in SiO_2 as written, they are intended simply to indicate that both molecular and atomic oxygen are assumed capable of reacting with the silicon to form SiO_2. Based upon their kinetic data, Van der Meulen and Ghez determined rate constants for the first reaction (k_1) and for the combination of reactions (1.28) and (1.29) (k_3). The corresponding activation energies were found to be 1.91 eV for k_1 and 0.58 eV for k_3. Because of this difference in activation energies, k_1

24

would be expected to dominate at higher temperatures and therefore a $p^{1.0}$ pressure dependence and the normal Deal-Grove relationship should apply at these temperatures. At lower temperatures, k_3 should dominate, leading to a $p^{0.5}$ pressure dependence and a growth law with different effective rate constants. Thus they concluded that oxidation by molecular oxygen was dominant at high T while the atomic oxygen reaction becomes increasingly important as T decreases.

Recently Blanc [1.45] has proposed a kinetic model for thin oxides which in essence uses only Eqs. (1.28) and (1.29). That is, he assumed that the reaction involving molecular oxygen [Eq. (1.27)] does not occur at all. Based upon this, he derived a growth law given by

$$\frac{1}{2} u + \frac{1}{4} (\exp 2u - 1) = bt \qquad (1.30)$$

where $ax = \sinh u$ and a and b are rate constants. For thin oxides, this predicts a linear parabolic growth law similar to [1.1], but with different effective rate constants. Blanc has compared his theoretical predictions with the data of Irene [1.46] at temperatures of 780°-980°C with apparently good agreement. As Blanc points out, however, his model predicts a $p^{0.5}$ pressure dependence for the interface reaction rate; this would seem to be in disagreement with the data of Van der Meulen and Ghez.

Other models have been proposed to account for the enhanced growth rates of thin oxide layers. Grove [1.47], for example, proposed that molecular oxygen is ionized at the outer SiO_2 surface, forming O_2^-. Coupled diffusion between the O_2^- and holes then effectively increases the O_2^- diffusion rate until an oxide thickness on the order of the extrinsic Debye Length is reached (~120 Å for dry O_2 and ~5 Å for H_2O). This model might also be expected to be at least partially successful in explaining the

experimental results of Jorgensen [1.48] in which externally applied electric fields were found to affect oxidation rates.

Accurate modeling of the thin oxide regime is most hindered today by the lack of a comprehensive body of experimental data. If such data did exist, it would then likely be possible to differentiate between the various proposed models or to define a new model. A major effort in this program has, therefore, been aimed at obtaining such a body of data. Since an oxidation rate law may result from a number of possible mechanisms, the observed reaction rate law may not be conclusive evidence for a particular mechanism. Therefore, a large matrix of experimental conditions has been sought in order to reduce the number of possible mechanisms involved by discarding those which are valid only under some of the exprimental conditions. Thus, the effects of temperature, partial pressure of oxygen in the oxidizing ambient, silicon orientation and the doping density of the silicon substrate were investigated. Oxidations were carried out in collaboration with E. A. Irene of IBM, using their in-situ ellipsometry capability and cover the temperature range of 800-1000°C, partial pressures of 0.01-1.0 atmospheres O_2, three substrate orientations and a variety of substrate doping levels.

The results which have been obtained (of which Fig. 1.10 is an example), suggest an enhancement of the oxidation rate of the form

$$\frac{dx}{dt} = \frac{B}{2X + A} + K_1 e^{-t/\tau_1} + K_2 e^{-t/\tau_2} \tag{1.31}$$

where the first term is the usual linear parabolic relationship and the second two terms correspond to decaying exponentials which dominate over the first ≈ 20 A and 200 A of growth respectively. A more detailed discussion of these results is contained in section 5 later in this report. SUPREM III contains a simplified version of eqn. 1.31 as indicated in Table 1.1. This
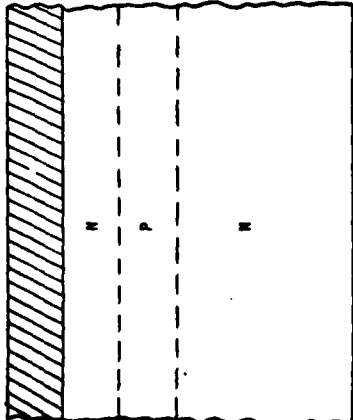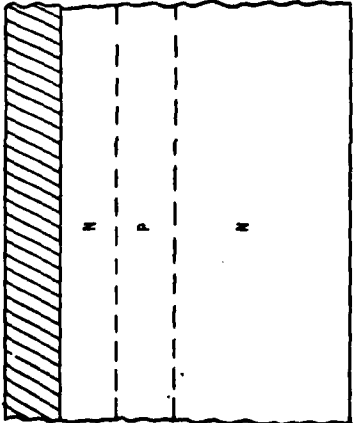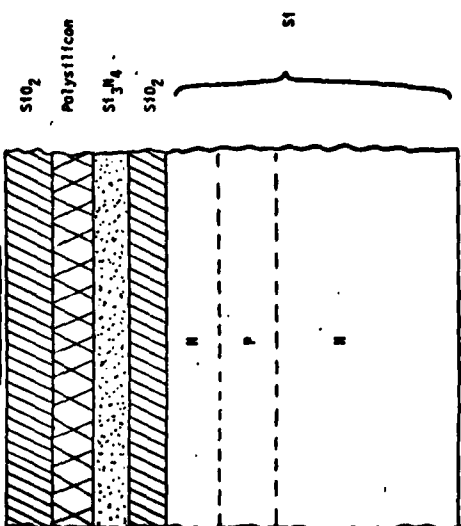
should greatly improve the predictive capability over SUPREM II (which artificially multiplied $(B/A)^i$ by 10X for X < 200Å) and SUPREM I which did not model the thin regime at all. At the present time we are still uncertain about the physical mechanisms responsible for eqn. (1.31).

2. ION IMPLANTATION

Past versions of SUPREM (Versions I and II) utilized essentially look-up tables to calculate range distributions as shown in summary form in Table 1.4. As device geometries shrink, the limitations of this approach become apparent, particularly because most modern device structures employ implants through masking layers. In addition to complicating calculated range distributions, such multi-layer structures introduce additional effects such as the possibility of knock-on ions from the masking layer. We first review here the method by which implanted profiles are calculated in SUPREM I and II and then describe a new approach, developed under this program and employing the Boltzmann Transport model, which is employed in SUPREM III. This new approach promises a much more general and robust process modeling capability.

The simplest description of an implanted impurity profile in silicon or silicon dioxide is a symmetrical Gaussian curve with first two moments the projected range R_p, and the standard deviation σ_p, calculated from the LSS theory [1.49]. However, experimental distributions of many ions, such as boron or arsenic, are found to be asymmetrical. The simple Gaussian approximation of those implanted profiles is often inadequate so that higher order moments must be used to construct range distributions. Gibbons and Mylroie [1.50] have shown that the third central moment is enough to provide sufficient information to construct accurate distributions when the asymmetry is not excessive (less than the standard deviation). In these cases, the distribution can be represented by two half-Gaussian profiles, each with a different standard deviation, σ_1 and σ_2, joined together at a model range Rm as

27

Table 1.4 - ION IMPLANTATION MODELS

SUPREM 1

2 layers
1 interface
2 materials

B, P, As, Sb:

Table look-up of R_p, σ_1, σ_2 for two half Gaussians joined at R_p.

SUPREM II

2 layers
1 interface
2 materials

P, As, Sb:

Table look-up of R_p, σ_1, σ_2 for two half Gaussians joined at R_p.

B:

Table look-up, Pearson IV distribution, empirical exponential tail added

SUPREM III

5 layers
4 interfaces
Any materials

B, P, As, Sb:

a) Boltzmann transport equation solution for primary ion range statistics through multi-layer targets.

b) Gaussian, two-sided Gaussian, and Pearson IV Distributions from table look-up at user option

28

shown in Fig. 1.11. This method can be used for profiles such as arsenic and phosphorus. However, for boron a modified Pearson IV distribution [1.51] is found to be more realistic. For the joint half-Gaussian distribution, the two sides are given by

$$C_1(y) = P_x \exp\left[-(y - R_m)^2/2\sigma_1^2\right] \qquad 0 \leq y \leq R_m \qquad (1.32)$$

$$C_2(y) = P_x \exp\left[-(y - R_m)^2/2\sigma_2^2\right] \qquad R_m < y \leq \infty$$

The parameters are determined from table look-up and interpolation [1.50, 1.52]. All ions (B, P, As, and Sb), are modeled this way in SUPREM I.

Hofker et al. [1.53] have shown that the implanted boron profiles in amorphous silicon before annealing, may be described by a Pearson-type IV distribution. They have experimentally determined the fourth moment ratio. Although this distribution describes well the profile of boron from the surface to a distance somewhat beyond the peak of the distribution, Fig. 1.12 shows that the approximation fails to model the observed exponential tail that is due to random scattering of boron ions along channeling directions. Based on experimental results [1.54-1.56] an empirically modified Pearson IV distribution can be used by adding an exponential tail with a fixed characteristic length (0.045 μm), independent of dose, energy and crystalline surface orientation [1.51]. The tail is attached to the shoulder of the standard Pearson IV distribution in SUPREM II, where the concentration drops to 50 percent of the peak value. Of course, after the addition of the tail, renormalization of the distribution to the implanted ion dose is necessary. Typical resulting profiles from this modification are shown in Fig. 1.12. These profiles are representative of those generated by SUPREM II.

Practical device processing very frequently involves implantation into target structures composed of a substrate on which one or more thin films of

29

varying composition have been grown or deposited. LSS theory assumes an iso-tropic or amorphous semi-infinite target. The effect on ion range of implan-tation into such a multilayer target is not considered. Also not modeled is the experimentally observed dopant profile's deep channeling tail that results from ions scattered into the relatively unobstructed open "channeling" direc-tions of the crystalline silicon substrate. Further, the secondary implanta-tion of "knock-on" atoms from a masking surface layer into the substrate as recoils from collisions with the primary ion beam critically affect electrical properties and device performance. This "recoil implantation" process also cannot be analyzed by LSS theory.

The substrate itself can be severely damaged by implantation. An ini-tially crystalline substrate may be driven amorphous or large densities of vacancy-interstitial point defect pairs created. Such damage distributions are integrally involved in implant annealing and activation, but have not been adequately modeled. An approach similar to that of LSS has been used by Winterbon et al. [1.57] to calculate energy deposition (damage) profiles. In this case, however, the asymmetry of such profiles necessitates the use of many moments for an accurate description. In addition, non-nuclear effects including electronic stopping must be neglected. Brice [1.58] has circumvent-ed these problems to some extent by using a variation of the LSS moments approach to determine the more symmetric intermediate range distributions and then calculates energy deposition profiles directly from these. Unfortunately such approaches again require the assumption of a homogeneous target and therefore fail to address the problem of multilayered media.

In a transport equation approach [1.59,1.60] quantities of interest are determined by calculating in a step-wise fashion the momentum distributions of both primary and recoil particles as a function of distance z from the target

surface. It is assumed that the target is amorphous or aligned in a random direction and possesses translational symmetry in all directions parallel to its surface. Thus only two components of momentum, or equivalently an energy and direction cosine, are necessary to define a particle's state of motion. Let F(p,z)dp be the number of particles (integrated over time) with momentum in the region dp about p and depth z. There will be one such distribution function for each projectile type. The spatial evolution of each distribution is governed by the Boltzmann transport equation:

$$\frac{\partial F(p)}{\partial} = N \int \left[\frac{F(\vec{p}^1)\, d\sigma\, (\vec{p}^1\vec{p})}{cox\ \theta_{\vec{p}}^1} - \frac{F(p)d\sigma(\vec{p}\to\vec{p}^1)}{cos\ \theta_{\vec{p}}} \right] + Q(\vec{p}) \tag{1.33}$$

Here for simplicity we assume a single element target with number density N. The total differential scattering cross section $d\sigma(\vec{p}\to\vec{p}^1)$ represents a differential area presented to an incident ion by each target atom. Any ion which enters this area with momentum p is scattered to momentun $\vec{p}^1$. A particle which scatters to a momentun (or energy) below some fixed limit is considered stopped and is removed from the momentum distribution. In this way the range profiles of particles are generated. The factors of cos (θ) result from the fact that a particle moving at an angle θ with respect to the z axis travels a distance dz/cos (θ) as the entire distribution moves the distance dz. The quantity Q is a generation term which (for recoil distributions) allows particles to be created from rest.

The distribution functons are assumed known at the surface plane z = 0. Recoil distributions are identically zero there and the momentum distribution of the primary ion is taken to be a delta function:

$$F(\vec{p},0) = \Phi\delta\ (\vec{p} - \vec{p}_0) \tag{1.34}$$

31

where ϕ is the total dose (per cm^2) and $\vec{p}$ is the (unique) momentum of ions in the incident beam. With these initial conditions the transport equation is integrated numerically to determine the distributions at all depths $z > 0$. Crossing a material interface during the integration, as occurs in the case of multilayered target, necessitates changing only the number density and cross section to quantities appropriate to the new material. No changes in the momentum distributions occur because such distributions are continuous functions of z for any target composition. Such an approach is thus easily applied to targets of practical interest.

To apply the transport equation (TE) effectively to problems in ion implantation some additional assumptions must be made. These assumptions are not inherent in the TE approach but arise from our limited knowledge of the scattering process itself and are typically made in all methods. In particular we assume (1) the scattering cross section $d\sigma$ is the sum of an elastic nuclear part $d\sigma_n$ and an inelastic electronic part $d\sigma_e$, (2) scattering consists of a series of independent binary events so that contributions from elements in a multi-element target may be simply summed, (3) electronic stopping acts only as an energy dependent drag force which removes energy continuously from the distribution, (4) only nuclear events generate recoils, produce lattice disorder, or cause significant angular deflections, (5) classical two body mechanics may be used to describe the angular deflections of both incident and recoil particles, and (6) for each nuclear event involving energy transfer T (assumed much greater than the binding energies of lattice atoms), a particle of energy T is created in the appropriate recoil distribution consistent with assumption 5.

The results of such TE calculations have been found to be quite sensitive to the differential nuclear cross section used. The scattering function

32

$f(t^{1/2})$ used in $d\sigma_n(t)$ is determined by parameter constants λ, m and q [1.61, 1.62]. The constants λ = 1.309, m = 0.333, and q = 0.667 were originally fit to an interatomic potential based on the Thomas-Fermi model of the atom and have been used in most work to date. However, Kalbitzer [1.63] has proposed that the parameters λ = 2.54, m = 0.25, and q = 0.475 will provide a better fit to a large collection of experimental data taken under conditions where nuclear stopping dominates (i.e., when the LSS-normalized energy $\varepsilon < 1$). The two scattering functions differ most in the low energy region and are compared in Fig. 1.13. We believe that the Kalbitzer cross section is better at low energy transfers but that near $\varepsilon = 1$ values closer to the Thomas-Fermi cross section should be used. We have therefore constructed a hybrid cross-section which agrees with that of Kalbitzer at low energy transfers but evolves continuously into the Thomas-Fermi cross section at higher transfers. The mark shown in Fig. 1.13 gives the approximate point where the hybrid crosses over from Kalbitzer to Thomas-Fermi. A cross section which is very similar to this hybrid but which has a more fundamental basis was recently proposed by Wilson, et al. [1.64].

In Fig. 1.14 TE calculations using both the Kalbitzer and hybrid cross sections are compared with a Pearson type IV distribution generated from LSS moments for a dose of 10^{16} arsenic per cm^2 at 355 keV. Also shown is experimental data from Hirao, et al. [1.65] of arsenic implanted into single crystal silicon substrates which were misaligned by 8° from the (111) direction. For direct comparison the Kalbitzer cross section was used in the LSS calculation. Under these conditions ($\varepsilon \simeq 0.5$) we expect nuclear scattering to be the dominant stopping mechanism and hence sensitivity to the nuclear cross section is most effectively examined. The figure shows that the TE and LSS calculations using the Kalbitzer cross section agree well near the peak of the distribution. However the LSS profile is more skewed, and both the TE and LSS

33

calculations yield profiles that are somewhat deeper than the experimental data would suggest. By comparison, the TE calculation using the hybrid cross section is seen to agree quite well with experiment. The disagreement which occurs in the deepest part of the distribution is most likely due to channeling effects.

Similar types of profiles can be easily obtained in multi-layer targets such as those illustrated in Table 1.4. This capability is implemented in SUPREM III. As an example, we consider an implantation of arsenic into a target consisting of a silicon substrate coated with silicon dioxide (SiO_2). During the implantation, oxygen atoms will be recoil-implanted from the oxide to the silicon. Since oxygen can introduce deep levels in silicon it is of interest to determine the total number of oxygen atoms (per cm^2) which ultimately reside in the silicon substrate. For a given initial ion energy, this number will be a function of the oxide thickness. Goezberger, et al. [1.66] have used surface state techniques to measure the dependence of total oxygen recoil flux on oxide thickness for an implant of 5×10^{13} cm^{-2} arsenic at 100 keV. We have calculated the same quantity using the Kalbitzer cross section with the results given in Fig. 1.15. As can be seen from the figure, the TE results are in good agreement within the experimental error. Such a calculation can be very valuable in assessing the merits of conventional implantation processes.

SUPREM III contains the full multi-layer modeling capability for the primary implanted ion (as shown in Table 1.4). It does not contain the capability to predict recoil profiles or damage distributions even though this is inherent in the Boltzmann Transport approach, because present general algorithms for implementing these capabilities require unacceptable amounts of computer time. A specific objective of future work is to reduce the required time and hence to add

these capabilities to the program. Some progress in this regard has been made in the specific case of O recoils from SiO_2 layers as described in our annual reports [1.15,1.16].

A final example of the new ion implantation modeling capability is shown in Figs. 1.16 and 1.17. Fig. 1.16 contains the SUPREM II and SUPREM III input command files for a bipolar transistor with an implanted base and emitter. This particular device is a high frequency structure with very shallow junctions. The outputs produced by SUPREM II and SUPREM III are shown in Fig. 1.17.

The arsenic emitter profiles are quite similar, both having the characteristic plateau shape and a junction depth of 0.18 microns. The most noticeable difference in the results of the two programs is the boron base region. SUPREM II models the boron base implant with a Pearson IV distribution having an empirically determined exponential tail added to the profile to account for residual channeling. By contrast SUPREM III uses the more fundamental Boltzmann transport approach to modeling the boron implantation, resulting in a narrower base region which much more closely fits the experimental data.

1.3. DIFFUSION

Impurity diffusion is an essential part of all semiconductor device fabrication. When all impurities are present at concentrations below the intrinsic carrier concentration at the process temperature, modeling of impurity diffusion is relatively straightforward. This is not the case, however, in many processing techniques of practical importance. In past versions of SUPREM (Versions I and II), we have generally relied on diffusion models developed elsewhere. Under this contract, specific experimental work has concentrated on "anomalous" diffusion which occurs during thermal oxidation of the silicon surface. New models, which incorporate such effects are included in

35

Table 1.5 - DIFFUSION MODELS

SUPREM 1
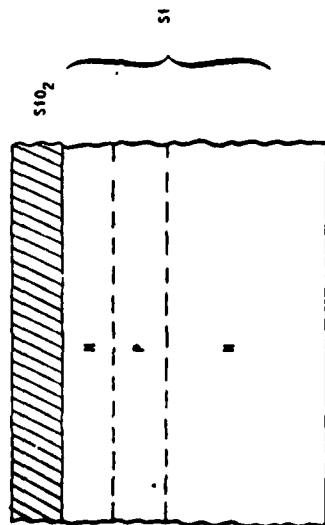
a) B: $D_B = D^X + D^+$

$$= D_i \left[\frac{1 + \beta \frac{n_i}{n}}{1 + \beta} \right]$$

$(\beta \approx 3)$

b) A, Sb, P: $D = D^X + D^-$

$$= D_i \left[\frac{1 + \beta \frac{n}{n_i}}{1 + \beta} \right]$$

$(\beta \approx 100)$

SUPREM II
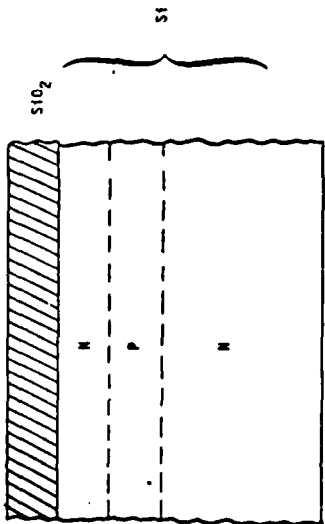
a) B: $D_B = f_e [D^X + D^+] + D_{OED}$

$$= f_e D_i \left[\frac{1 + \beta \frac{n_i}{n}}{1 + \beta} \right] + D_{OED}$$

$(\beta \approx 3)$

b) As: $D_{AS} = f_e [D^X + D^-]$

$$= f_e D_i \left[\frac{1 + \beta \frac{n}{n_i}}{1 + \beta} \right]$$

$(\beta \approx 100)$

c) Sb: $D_{Sb} = f_e [D^X + D^-]$

$$= f_e D_i \left[\frac{1 + \beta \frac{n}{n_i}}{1 + \beta} \right]$$

$(\beta \approx 1)$

SUPREM III
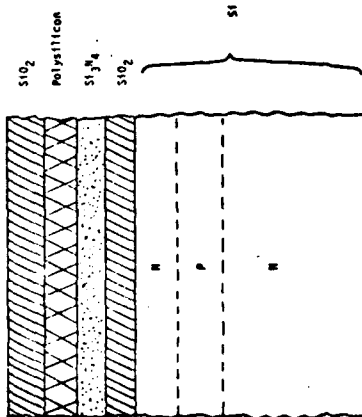
a) B: $D_B = D^X + D^+ + D_{OED}$

$$= D_i^X + \frac{p}{n_i} D_i^+ + D_{OED}$$

b) As: $D_{AS} = D^X + D^- + D_{OED}$

$$= D_i^X + \frac{n}{n_i} D_i^- + D_{OED}$$

c) Sb: $D_{Sb} = D^X + D^- + D_{OED}$

$$= D_i^X + \frac{n}{n_i} D_i^- + D_{OED}$$

36

SUPREM III

d) P: $D_p = D^x + D^- + D^= + D_{OED}$

$= D_i^x + \dfrac{n}{n_i} D_i^- + \left(\dfrac{n}{n_i}\right)^2 D_i^= + D_{OED}$

(surface region)

$D_p = D_i^x + D_i^- \left\{ \dfrac{n_s^3}{n_e^2 n_i}\, e^{3\Delta Eg/kT} \left[\dfrac{0.3eV}{1+e^{\frac{}{kT}}}\right] \left[e^{-\frac{x-x_p}{25\,\mu}}\right]\right\} + D_{OED}$

(Tail region)

e) OED: $D = D_{N_2} + \Delta D_{OX}$

$\Delta D_{OX} = f_{II}\, D_{Si}^i \left(\dfrac{Cv^i}{Cv}\right) \left[\dfrac{K_1\, e^{-X/L_D}}{P\sqrt{1 + K_2\, P_{HCl}^{0.5}}}\right] \left(\dfrac{dx}{dt}\right)^{0.5}$

where $F_{II} = K\, e^{-EA/kT}$ and varies with impurity

SUPREM II

d) P: $D_p = f_e\left[D^x + D^=\right] + D_{OED}$

$= f_e \left[D_i^x + \left(\dfrac{n}{n_i}\right)^2 D_i^=\right]$

(Surface region)

$D_p = f_e \left\{D_i^x + D_i^- \dfrac{n_s^3}{n_e^2 n_i}\left[1 + e^{0.3\ eV/kT}\right]\right\} + D_{OED}$

(Tail region)

e) OED: B: $D_B = D_{B_{N_2}} + K e^{-EA/kT}$

P: $D_{P_{O_2}} = 1.8\, D_{PN_2}$

$D_{P_{H_2O}} = 3.3\, D_{PN_2}$

As, Sb: None

SUPREM III. In the paragraphs below we outline first the models used for diffusion in non-oxidizing ambients and then review some of the recent work under this program which addresses the additional effects which occur in oxidizing ambients. Table 1.5 summarizes the diffusion models which will be discussed.

The most general case of impurity diffusion in silicon may involve a non-constant diffusivity along the impurity concentration gradient, as in high concentration diffusions in which the impurity concentration exceeds the intrinsic carrier concentration $n_i(T)$ at process temperature. An effective electric field driving force or electrostatic potential gradient may be established by the concentration gradient of either the impurity under question or another impurity present at greater concentration. Further, the long-range effects induced by nonequilibrium phenomena such as oxidation of the surface alter diffusion deep into the bulk. Simulation of such processes must involve the complete one-dimensional continuity equation

$$\frac{\partial C}{\partial t} = \frac{\partial}{\partial X}\left(D\,\frac{\partial C}{\partial X}\right) \pm \frac{q}{kT}\,\frac{\partial}{\partial X}\left(D\hat{C}\,\frac{\partial \phi}{\partial X}\right) \qquad (1.35)$$

where D is the effective diffusivity, and C and $\hat{C}$ are the total and electrically charged impurity concentrations, respectively. The potential ϕ is

$$\phi = \frac{kT}{q}\,\ell n\,\frac{n}{n_i} \qquad (1.36)$$

where n and n_i are the electron and intrinsic carrier concentrations, respectively. Thus, the first term in the continuity equation represents classical concentration gradient-driven diffusion, including non-constant diffusivity. The second term incorporates the electrostatic potential gradient-driven flux and also includes non-constant diffusivity. The sign of this field-driven flux term depends upon the charge polarity of the field-driven species.

38

Physically rigorous simulation of a diffusing impurity should therefore track all of the different mobile species, both charged and neutral, with which the impurity atom may associate. Further, diffusivities of each of the species must be known and possible interactions and/or exchanges among the different species may be significant. Physical understanding of such processes is incomplete, however. The numerical tasks presented by such complete simulation are also formidable.

Rather, the diffusion models in SUPREM at present represent an attempt to utilize the more accurate point defect-assisted diffusion models developed both in the literature and in work under this program. Vacancy-assisted models have been the most well-developed, and include the widest range of process conditions, particularly under non-oxidizing surfaces. Interstitial-assisted diffusion has been used to complement these models to simulate altered diffusivities under oxidizing surfaces.

Under non-oxidizing conditions, group III and V impurity atoms in silicon have been modeled to diffuse largely by silicon vacancy-assisted mechanisms, such as vacancy-impurity pair diffusion. Thus the diffusion coefficient is proportional to the concentration of such point defects. Although the concentration of neutral defects at any given temperature is independent of the impurity concentration (so long as it does not approach that of silicon atoms), the concentrations of defects with various charge states (which have been identified within the silicon bandgap) depend on the Fermi level position in the bandgap and thus are a function of impurity concentration. Recently, Fair [1.67] summarized some of the latest work on diffusion in silicon and suggested that the effective diffusion coefficient should be the sum of several diffusivities, each accounting for impurity interactions with different charge states of lattice vacancies. The effective diffusion coefficient under non-oxidizing conditions D_N can thus be expressed as

39

$$D_N = D_i^X + D_i^- [V^-] + D_i^= [V^=] + D_i^+ [V^+] \qquad (1.37)$$

where D_i^v is the intrinsic diffusivity due to each identified charge state, v, of vacancies (v: =, -, x, +; i.e., doubly negative, singly negative, neutral, and positive), and $[V^v]$ is the concentration of vacancies in each charge state, normalized to the intrinsic concentration of that state. Using the Boltzmann approximation it can be easily shown that these normalized concentrations may be given by

$$[V^-] = \frac{n}{n_i} \quad , \quad [V^=] = \left(\frac{n}{n_i}\right)^2 \text{ and } [V^+] = \frac{n_i}{n} \qquad (1.38)$$

where n is the free electron concentration. Thus under intrinsic conditions ($n = n_i$), Eq. (1.37) becomes

$$D_i = D_i^X + D_i^- + D_i^= + D_i^+ \qquad (1.39)$$

i.e., the intrinsic diffusivity is the sum of the diffusivities resulting from the various vacancy charge states.

Often, not all of the intrinsic diffusivities due to the various vacancy charge states are significant. For each commonly used dopant, specific vacancy charge states have been found to be dominant in determining D_N while the contributions of other charge states are less significant and may be neglected.

1.3.A. Boron

Boron as an acceptor is negatively charged in the silicon lattice. It diffuses primarily with V^+ and V^X vacancies. Thus, boron diffusivity is largely determined by the D^+ and D^X contributions [1.68,1.69]:

$$D_N (B) = D_i^X + D_i^+ \left(\frac{p}{n_i}\right)$$

$$= \left[0.037 + 0.72 \left(\frac{p}{n_i}\right)\right] \exp \left(\frac{-3.46 \text{ eV}}{kT}\right) \quad cm^2/sec \qquad (1.40)$$

In this instance, V^+ and V^X assisted diffusion of the dopant exhibit similar activation energies. It may be seen that D^+ in particular will dominate for boron except in heavily doped n-type silicon.

SUPREM III implements eqn. (1.40) for Boron intrinsic diffusivity. Earlier versions of SUPREM (I and II) utilized approximate versions of this expression due to the lack of detailed information at the time those codes were written. In SUPREM I and II, the form of the diffusion coefficient implemented is given by

$$D_B = D_i \left[\frac{1 + \beta\left(\frac{p}{n_i}\right)}{1 + \beta}\right] \qquad (1.41)$$

where D_i is the intrinsic diffusivity, which by comparison with Eqn. (1.40), contains contributions from neutral and + charged vacancies. $\beta = D^+/D^X$ is a measure of the effectiveness of charged vacancies relative to neutral ones and was empirically determined to be $\simeq 3$ for Boron. Table 1.5 summarizes these results.

1.3.B. Arsenic

Arsenic as a donor appears to diffuse primarily with V^X and V^- vacancies. The diffusivity becomes [1.69]:

$$D_N (As) = D_i^X + D_i^- \left(\frac{n}{n_i}\right) \qquad (1.42)$$

$$= 0.066 \ \exp\left(\frac{-3.44 \text{ eV}}{kT}\right) + 12.0 \ \left(\frac{n}{n_i}\right) \ \exp \left(\frac{-4.05 \text{ eV}}{kT}\right) \ cm^2/sec$$

41

In this instance D_i^X and D_i^- are comparable throughout the process temperature range. The relative significance of D^X vs. D^- diffusion of arsenic will then depend entirely upon the doping level and "polarity" of the silicon. SUPREM III implements eqn. (1.42) for As diffusion. Earlier versions (I and II) used an approximate form (eqn. 1.41) with $\beta \simeq 100$. This value of β was empirically determined to provide reasonable agreement between experiment and theory.

At moderate doping concentrations, the net electron concentrations may be taken to equal the As concentrations, and application of eq. (1.42) is straightfoward. With high doping levels, however, the electron concentration n will deviate significantly fom the total arsenic C_T. Further, this n vs. C_T dependence may depend critically on the nature of the As doping source.

For ion implanted arsenic, the total arsenic concentration is [1.69]

$$C_{T_{II}}(As) = C_{(As^+)} + C_{(As^+)(V^-)} + C_{(As^+)(V^=)} \qquad (1.43)$$

To satisfy charge neutrality requirements, the net electron concentration becomes

$$n = C_{(As^+)} - C_{(As^+)(V^=)} \qquad (1.44)$$

With high doping levels the latter two species in eqn. (1.43) will account for substantial fractions of the total As, and n will not increase directly with C_T. Mass-action relates each of the constituent As species in eqn. (1.43) to n. The $(As^+)(V^-)$ pair is found to be relatively less significant in contributing to C_T. Thus, with

$$C_{(As^+)} \propto n$$

$$C_{(As^+)(V^=)} \propto n^3 \qquad (1.45)$$

42

the dependence of n vs. C_T resulting from eqn. (1.43) is well approximated by

$$C_{T_{II}} (\text{As}) = n + K(T) \, n^3 \qquad (1.46)$$

where $K(T)$ exhibits a slight temperature dependence

$$K(T) \simeq 9.41 \times 10^{-44} \exp\!\left(\frac{+0.33 \text{ eV}}{kT}\right) \qquad \text{cm}^6 \qquad (1.47)$$

For chemical source arsenic diffusions, the deviation of n from C_T at high doping levels appears even greater. Arsenic clustering models have been invoked to rationalize this dependence. Fair and Weber [1.69,1.70] have proposed that, in addition to the $(\text{As}^+)(\text{V}^-)$ and $(\text{As}^+)(\text{V}^=)$ pairs formed, another complex appears

$$2 \text{ As}^+ + \text{V}^= = (\text{As}^+)_2 \, (\text{V}^=) \qquad (1.48)$$
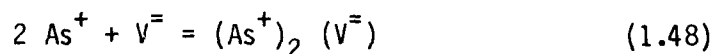
Two As atoms are involved via the reduction of As_2O_3 molecules in the chemical source by the Si surface. Vapor phase doping likely involves As_2 and As_4 species, while again monatomic As is negligible [1.69]. Formation of the $(\text{As}^+)_2(\text{V}^=)$ complex becomes much more significant than with implanted As diffusion. Thus eqn. (1.43) becomes

$$C_{T_{\text{Chem}}} = C_{(\text{As}^+)} + C_{(\text{As}^+)(\text{V}^-)} + C_{(\text{As}^+)(\text{V}^=)} + C_{(\text{As}^+)_2(\text{V}^=)} \qquad (1.49)$$

Here, the $(\text{As}^+)_2(\text{V}^=)$ species appears more significant in contributing to C_T than $(\text{As}^+)(\text{V}^-)$ or $(\text{As}^+)(\text{V}^=)$. With mass-action requirements, the n vs. C_T dependence is found to be well approximated in the case of chemical sources by

$$C_{T_{\text{Chem}}} (\text{As}) \simeq n + K'(T) \, n^4 \qquad (1.50)$$

43

Recent data of Murota, et al. [1.71] have confirmed this n^4 dependence. The temperature dependence of a two-As atom cluster contribution to C_T has been estimated to be ≈ 1.6 eV by Schwenker, et. al [1.72]. Thus,

$$K'(T) \approx 1.13 \times 10^{-68} \exp\left(\frac{+1.6 \text{ eV}}{kT}\right) \text{ cm}^9 \tag{1.51}$$

1.3.C. Phosphorus

Phosphorus has been observed to exhibit perhaps the most complex diffusive behavior of the commonly used dopants. Particularly in high concentration diffusions as for emitter regions of bipolar devices, such "anomalous" features as both enhanced and retarded diffusion, "kink" formation, enhanced "tail" diffusivity, and emitter-base push (i.e., enhancement of the diffusivity of other impurities such as boron present in the silicon beneath the high phosphorus region) have been observed. No single unified physical treatment of phosphorus diffusion has satisfactorily reconciled all of this anomalous behavior. However, models, particularly from the work of Fair [1.69,1.73], have been developed to allow reasonable simulation of many of these features of the diffused phosphorus profile.

1.3.C.1. Isoconcentration Diffusion

In the absence of a rapidly changing profile gradient, the basic phosphorus diffusion model is a logical extension of the mechanisms of impurity diffusion discussed above. Phosphorus as a donor diffuses with V^X, V^-, and $V^=$ vacancies [1.69]

$$D_N(P) = D_i^X + D_i^-\left(\frac{n}{n_i}\right) + D_i^=\left(\frac{n}{n_i}\right)^2$$

$$= 3.85 \exp\left(\frac{-3.66 \text{ eV}}{kT}\right) + 4.44\left(\frac{n}{n_i}\right)\exp\left(\frac{-4.00 \text{ eV}}{kT}\right)$$

$$+ 44.2\left(\frac{n}{n_i}\right)^2 \exp\left(\frac{-4.37 \text{ eV}}{kT}\right) \text{ cm}^2/\text{sec} \tag{1.52}$$

44

For lightly doped silicon, D^X dominates the diffusive behavior. With heavily n-type material, the charged vacancy pair-diffusion, in particular $D^=$, will largely determine the phosphorus diffusivity, because of the $(n/n_i)^2$ dependence. SUPREM I again implemented a simplified version of eqn. (1.52) (i.e. eqn. 41 with p/n_i replaced by n/n_i and $\beta \simeq 100$). SUPREM II used in essence eqn. (1.52).

As with arsenic, the electron concentration n will deviate significantly from the total phosphorus C_T at high doping levels. In analogy with eqs. (1.43) and (1.44), the total phosphorus and electron concentrations are [1.69,1.73]:

$$C_T(P) = C_{(P^+)} + C_{(P^+)(V^-)} + C_{(P^+)(V^=)} \tag{1.53}$$

$$n = C_{(P^+)} - C_{(P^+)(V^=)} \tag{1.54}$$

The resulting n vs. C_T dependence is well approximated by

$$C_T(P) = n + K''(T)n^3 \tag{1.55}$$

$$K''(T) \simeq 5.33 \times 10^{-43} \quad \exp\left(\frac{+0.40 \text{ eV}}{kT}\right) \quad cm^6 \tag{1.56}$$

1.3.C.2.   High-Concentration Profiles

With high-concentration surface-diffused layers of phosphorus, many more features must be taken into consideration. A typical profile is composed of three regions as shown schematically in Fig. 1.18.

a.   Surface Region

In the surface region, the diffusivity is given by Eq. (1.52) and may be seen to be dominated by $D^=$. The diffusivity will increase with increasing surface concentration, until at concentrations above ~3-4 $\times 10^{20}$ cm^{-3} the diffusivity will then begin to decrease. This effect may be attri-

buted to silicon bandgap narrowing due to the strain induced on the lattice by the combination of the finite phosphorus misfit and the relatively high solubility of phosphorus i the silicon lattice. This misfit strain-induced bandgap narrowing may be surface orientation dependent, and is estimated by [1.74,1.75,1.76]:

$$\Delta Eg_{mf} \langle 111 \rangle = -1.5 \times 10^{-22} \; [C_{T_S} - C_D(T)] \; eV \qquad (1.57)$$

$$\text{where} \quad C_D(T) = \frac{2.43 \times 10^{26}}{[T(°C)]^2}$$

$$\Delta Eg_{mf} \langle 100 \rangle = 1.23 \; \Delta Eg_{mf} \langle 111 \rangle \qquad (1.58)$$

where C_{T_S} is the total surface concentration of phosphorus (neglecting the interface pileup within ~100 Å of the Si/SiO$_2$ interface).

Another strain effect which must also be included affects the diffusivity of ion-implanted phosphorus profiles. In this case, it is observed that the diffusivity is further decreased under high dose conditions, probably due to permanent lattice disorder [1.74,1.76] produced by the implantation process. Thus in the case of ion-implanted phosphorus, an additional bandgap narrowing is calculated from [1.76]

$$\Delta E_{g_{II}} = -2.3 \times 10^{-6} \; (Q_p)^{0.25} \; eV \qquad (1.59)$$

where Q_p is the phosphorus atom dose (cm^{-2}).

Thus, the total bandgap narrowing

$$\Delta E_g = \Delta E_{g_{mf}} + \Delta E_{g_{II}} \qquad (1.60)$$

may be quite substantial, and must be accounted for by replacing n_i in eqn. (1.52) with an increased effective intrinsic carrier concentration

$$n_{ie} = n_i \; exp\left(\frac{-\Delta E_g}{2kT}\right) \; cm^{-3} \qquad (1.61)$$
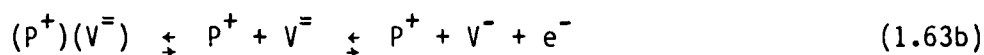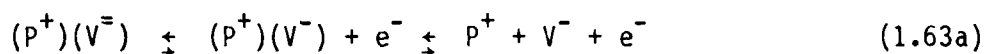
The importance of including the bandgap narrowing effect is illustrated by Fig. 1.19 where the calculated diffused phosphorus profiles with and without ΔE_g are compared.

With these considerations, the surface region diffusivity may be calculated and is found to be monitonically decreasing with depth into the silicon (Fig. 1.18), until the end of this region occurs at the point where the Fermi level drops below 0.11 eV from the conduction band edge, crossing the ionization level identified with $V^=$ vacancies [1.77]; a characteristic electron concentration n_e is defined at this Fermi level value

$$n_e \simeq 4.65 \times 10^{21} \exp\left(\frac{-0.39 \text{ eV}}{kT}\right) \text{ cm}^{-3} \qquad (1.62)$$

b. Tail Region

The key assumptions of this modeling are now invoked. Diffusion of phosphorus in the surface region has been largely via $(P^+)(V^=)$ pairs created at the surface. When the Fermi level crosses and falls below the $V^=$ ionization level at the characteristic n_e, $V^=$ vacancies whether "free" or paired with the ionized P^+ will lose an electron, with V^- vacancies resulting

$$(P^+)(V^=) \; \underset{\rightarrow}{\leftarrow} \; (P^+)(V^-) + e^- \; \underset{\rightarrow}{\leftarrow} \; P^+ + V^- + e^- \qquad (1.63a)$$

$$(P^+)(V^=) \; \underset{\rightarrow}{\leftarrow} \; P^+ + V^= \; \underset{\rightarrow}{\leftarrow} \; P^+ + V^- + e^- \qquad (1.63b)$$

Because the binding energy of the $(P^+)(V^-)$ pair is ~0.3 eV lower than that of the $(P^+)(V^=)$ pair, the dissociation of phosphorus-vacancy pairs and creation of free V^- vacancies is more likely and proceeds more rapidly via reaction (1.63a) than (1.63b). Both reactions will contribute to creating a supersaturation of V^- vacancies above and effectively independent of the equilibrium concentration corresponding to the local doping or Fermi level.

47

This supersaturation, as implied by eqn. (1.63), will be correlated with the concentration of $(P^+)(V^=)$ pairs originally created at the surface. The excess vacancies have a finite but substantial lifetime or diffusion length estimated to be ~1/4 sec or ~25 μ at temperatures in the range ~900 -1100°C [1.73,1.78]. Thus, they will enhance phosphorus diffusivity in the "tail" region substantially above that called for by the local Fermi level; and, this enhancement will be virtually constant, decaying only slowly deep into the bulk.

By applying these assumptions and invoking mass-action, the V^- vacancy supersaturation and enhanced tail diffusivity are estimated (with bandgap narrowing effects discussed above explicitly included) [1.73,1.74]:

$$D_{TAIL} = D_i^X + D_i^- \frac{[V^-]}{[V^-]_i}$$

(1.64)

$$= D_i^X + D_i^- \left\{ \frac{n_s^3}{n_e^2 n_i} \quad \exp\left(\frac{3\Delta E_g}{kT}\right)\left[1 + \exp\left(\frac{+0.3 \text{ eV}}{kT}\right)\right] \exp\left(\frac{-(X-X_e)}{25 \text{ μ}}\right)\right\}$$

where n_s is the surface electron concentration and X_e is the depth at which n_e is reached. Note that the decay length of this enhancement extends considerably beyond X_e and in fact well beyond the point at which the phosphorus doping reaches n_i and intrinsic diffusivity D_i would otherwise be reached. An enhancement factor G corresponding to the vacancy supersaturation driving this increased diffusivity may be defined for later use below:

$$G = \frac{D_i^X + D_i^- [V^-]/[V^-]_i}{D_i^X + D_i^-} \approx \left. \frac{D_{TAIL}}{D_i^X} \right|_{phosphorus}$$

(1.65)

c. Intermediate (Transition) Region

The transition from the diffusivity D_e at the end of the surface region at X_e to the enhanced diffusivity D_{TAIL} occurs over a finite depth that varies from ~1500 Å at 875°C to ~500 Å at 1000°C and vanishes for temperatures in excess of 1050°C. A "kink" in the phosphorus profile is observed at the point at which D_{TAIL} is reached. This transition region is perhaps the least well understood in the phosphorus profile. It is not unreasonable that the phosphorus-vacancy pair dissociation reactions of eqs. (1.63) driven by Fermi level changes along the profile gradient occur not instantaneously but over a finite depth.

Empirically, Boltzmann-Matano analysis has indicated that the diffusivity in this region <u>increases</u> with depth in proportion to n^{-2} [1.73, 1.79]. Thus, the diffusivity calculated at X_e using eqn. (1.52) is multiplied by $(n_e/n)^2$ with further depth until the point at which D_{TAIL} of eqn. (1.64) is reached and the tail region begins.

d. Emitter-Base Push

Because of the long decay length of the excess vacancies created by the phoshporus diffusion, the vacancy supersaturation will exist throughout the tail region and even past any metallurgical junction into which the phosphorus may diffuse. The charge state of the excess vacancies will likely continue to change in response to the local Fermi level. Thus, the vacancy supersaturation approximated by the enhancement factor G defined above will enhance the diffusion not only of phosphorus in the tail region, but also of other impurities of the dominant "polarity" type both in the tail region and beneath the tail region. For example, the empirically observed "emitter-base push" enhancement of boron diffusivity in the p-type base region beneath a phosphorus n-type emitter region becomes [1.73]

$$D_{BASE} (B) = D_i (B) [G] \qquad\qquad (1.66)$$

where D_i (B) is the intrinsic boron diffusivity in the base region.

1.3.D.   Antimony

Antimony as a donor diffuses with V^X and V^- vacancies [1.69],

$$D_N (Sb) = D_i^X + D_i^- \left(\frac{n}{n_i}\right)$$

$$= 0.214 \exp\frac{-3.65\ eV}{kT} + 15.0\left(\frac{n}{n_i}\right) \exp\frac{-4.08\ eV}{kT}\ cm^2/sec \qquad (1.67)$$

D_i^- is somewhat greater than D_i^X at process temperatures and will dominate in n-type material.   As with B, As and P, SUPREM I and II used a simplified expression for Sb diffusion, given by eqn. 1.41.

1.3.E.   Solid Solubility

The above models must be applied with care particularly as the dopant concentration approaches its solubility limit in silicon.   This limit establishes $C_{T_{max}}$ in applying eqns. (1.46) and (1.50) for arsenic, and eqn. (1.55) for phosphorus, for example.   Dopant in excess of $C_{T_{max}}$ will likely precipitate and is considered not diffusable, not contributing to the n vs. C_T dependence of arsenic and phosphorus, and not contributing to the misfit strain-induced bandgap narrowing $\Delta E_{g_{mf}}$ in the phosphorus model [1.74].

1.3.F.   Diffusion Under Oxidizing Surfaces

The oxidizing surface has long been observed to alter the diffusive behavior of the commonly used dopant impurities even at considerable depth into the bulk.   Boron, phosphorus, and arsenic exhibit oxidation-enhanced diffusion (OED) [1.80,1.81], as shown for example in Fig. 1.20.   Antimony, on the other hand, has recently been reported to demonstrate oxidation-retarded

diffusion (ORD) [1.82]. While some measure of understanding and quantitative modeling of OED has been achieved, ORD is a more recently observed phenomenon and presents new challenges to modeling.

As diffusion is most probably a silicon point defect-assisted process, consideration of OED and ORD must begin with an understanding of the influence of the oxidation process on the substrate silicon point-defect population kinetics. A reasonably current view of the oxidizing interface is shown in Fig. 1.3, which represents a composite of the ideas of many groups. Workers such as Sanders and Dobson [1.83] have for some time suggested that the oxidizing interface acts as a vacancy sink and interstitial source. Evidence for this view has been indirect but substantial. For example, interstitial generation during oxidation has formed the basis of successful modeling of OED and the related growth of oxidation-induced stacking faults (OISF) [1.84]. Vacancy consumption in the interface oxidation reaction has been proposed as the basis for the enhanced oxidation of heavily doped silicon [1.7,1.8]. If these reactions at the oxidizing interface do in fact occur, then their influence on impurity diffusion may be addressed.

1.3.F.1.   Dual Mechanism Impurity Diffusion

As has been noted, impurity diffusion via vacancy-assisted mechanisms has already yielded well-developed models, particularly for the non-oxidizing interface. Interstitial-assisted diffusion mechanisms, such as the Watkins relacement (kick-out or partial interstitialcy processes [1.75, 1.85] have been visualized, but models utilizing these mechanisms in silicon have been less well-developed. A general formulation for impurity diffusion would consider both possibilities, and propose both substitutional and interstitial impurity populations, C_{SI} and C_{II}, respectively, diffusing together through the Si lattice.

(The precise nature of an interstitial impurity, as with Si self-interstitials or even Si vacancies, may not be obvious; a "pure" interstitial with perhaps four dangling/broken bonds in the lattice would seem energetically to be unfavorable. The impurity atom may exhibit some partial interstitial nature, perhaps by sharing a lattice site with a silicon "interstitial." Thus dopant atoms of this nature may diffuse via interstitial-assisted mechanisms, while purely substitutional impurity atoms diffuse through vacancy-assisted processes. This allows the total impurity concentration to be separated into substitutional and interstitial populations.)

Thus, we may define the total impurity concentration C_T as

$$C_T = C_{SI} + C_{II} \qquad (1.68)$$

and overall dopant diffusion becomes [1.75]

$$\frac{\partial C_T}{\partial t} = D_{SI} \left(\frac{\partial^2 C_{SI}}{\partial X^2} \right) + D_{II} \left(\frac{\partial^2 C_{II}}{\partial X^2} \right) \qquad (1.69)$$

where D_{SI} and D_{II} are the vacancy-assisted and interstitial-assisted diffusivities, respectively.

In thermal equilibrium, under non-oxidizing conditions, and for lightly-doped silicon, the intrinsic interstitial impurity flux will be some fraction f_{II} of the intrinsic substitutional impurity flux

$$D_{II}^i \, C_{II}^o = f_{II} \, D_{SI}^i \, C_{SI}^o \qquad (1.70)$$

where f_{II} will vary with the impurity. It seems likely that f_{II} will also contain some temperature dependence if the diffusion processes via interstitials vs. vacancies exhibit different activation energies.

The defect and impurity populations are not independent, and exchanges may take place among the various species in the silicon lattice--vacancies, interstitials, substitutional impurities, and interstitial impurities. One likely process, the Watkins replacement mechanism, involves creation of interstitial impurities I_I (and substitutional silicon Si_S) via exchange between silicon interstitials Si_I and substitutional impurities I_S,

$$Si_I + I_S \rightleftharpoons Si_S + I_I \tag{1.71}$$

$$C_{II} = \frac{K_1(T) \, C_{SI} \, C_I}{C_{Si}} \approx K_1'(T) \, C_{SI} \, C_I \tag{1.72}$$

where C_I is the silicon interstitial concentration and $K_1'(T)$ is a temperature dependent reaction constant. The concentration of substitutional silicon C_{Si} is so large ($\sim$ the density of lattice sites) compared to the other species that it may be assumed unaffected by the reaction and absorbed into $K_1'(T)$.

With this exchange process, and substituting from eqn. (1.70), the interstitial impurity flux may be approximated under general non-equilibrium conditions through the substitutional impurity flux and the silicon interstitial population

$$D_{II}C_{II} \sim \left(\frac{D_{II}}{D_{II}^i}\right) f_{II} \, D_{SI}^i \, C_{SI}\left(\frac{C_I}{C_I^0}\right) \tag{1.73a}$$

Variation of the interstitial-assisted flux with total interstitial population is explicitly included. However, the dependence of (D_{II}/D_{II}^i) on doping is not obvious because of uncertainty concerning the possible charge states of silicon interstitials that may make different contributions, with different temperature and doping dependencies, to the interstitial assisted

53

diffusivity. With no additional information available, D_{II} will be assumed relatively unaffected by doping, and (1.73a) becomes

$$D_{II} C_{II} \approx f_{II} \ D_{SI}^{i} \ C_{SI} \left(\frac{C_I}{C_I^{\ 0}} \right) \qquad (1.73b)$$

If the Si interstitial population is assumed to vary relatively little with depth X into the silicon, then eqn. (1.69) becomes

$$\frac{\partial C_T}{\partial t} = \left[D_{SI} + f_{II} \ D_{SI}^{i} \left(\frac{C_I}{C_I^{\ 0}} \right) \right] \frac{\partial^2 C_{SI}}{\partial X^2} \qquad (1.74)$$

The commonly used dopants are assumed to be largely substitutional (diffusing mostly via vacancy-assisted mechanisms) in silicon. Eqn. (1.74) may then be considered the diffusion equation for the total impurity concentration with an effective diffusivity that combines a dominant vacancy-assisted diffusion process with a supplemental interstitial-assisted diffusion mechanism,

$$D = D_{SI} + f_{II} \ D_{SI}^{i} \left(\frac{C_I}{C_I^{\ 0}} \right) \qquad (1.75)$$

This effective diffusivity will vary with vacancy concentration through D_{SI}. Disturbances in the populations of silicon interstitials or vacancies will also alter the apparent overall diffusivity by driving exchanges between the two populations of impurities, each with different intrinsic diffusivities, to shift the relative proportion of impurities diffusing by vacancy vs. interstitial-assisted mechanisms. This exchange may occur via reaction (1.71), yielding the explicit interstitial dependence in eqn. (1.75). A second possible reaction may involve the combination of an interstitial impurity with a silicon vacancy to produce a substitutional impurity

$$Si_V + I_I \ \rightleftarrows \ I_S \qquad (1.76)$$

$$C_{SI} = K_2 (T) C_{II} C_V \qquad (1.77)$$

(Implied in reactions (1.71) and (1.76) is the annihilation of a silicon interstitial with a silion vacancy to reproduce substutional silicon,

$$Si_V + Si_I \updownarrow Si_S \qquad (1.78)$$

$$C_V C_I = K_3(T) C_{Si} \approx K_3' (T) \qquad (1.79)$$

This reaction implies a mass-action relation between the vacancy and interstitial concentrations that applies in thermal equilibrium. The significance of this mass-action in diffusion processes would be dependent on the activation energy barrier to the reaction, which will determine the time constant required to recover from non-equilibrium conditions.)

1.3.F.2.   Silicon Interstitial Generation

The influence of the oxidizing interface on impurity diffusivity will depend on disturbances produced in the point defect populations. Interstitial generation will depend on the oxidation and substrate parameters.

Fair [1.75] contends that the limiting reaction at the oxidizing interface that leads to interstitial generation depends upon interstitial oxygen ions $O_I^=$. The interstitial generation rate should exhibit a half-power dependence on oxidation rate $(dx_0/dt)^{0.5}$. (An alternative explanation for this power dependence is given in section 10 of this report.)

The interstitials created are subject to recapture at surface kink sites during surface regrowth [1.84]. The surface kink density ρ is dependent on surface orientation, with [1.75]

$$\rho \langle 111 \rangle \approx 1.5 \, \rho \langle 100 \rangle \qquad (1.80)$$

55

With increasing substrate doping, whether n or p type, the total equilibrium concentration of vacancies will increase [1.23]. The increasing vacancies will reduce the interstitial concentration by recombination reaction (1.78). This may be considered a manifestation of a mass-action relation previously considered. The time constant required is not known, but the possible mass-action with its activation energy barrier may in any event yield a direct proportionality

$$\frac{C_I}{C_I{}^i} \propto \frac{C_V{}^i}{C_V} \tag{1.81}$$

where C_V^i and C_I^i are the equilibrium concentrations in lightly-doped silicon.

The above annihilation of interstitials and time constant will be manifest as a finite decay length or lifetime of the generated interstitials. Thus, the generated interstitial concentration will decrease with depth X away from the oxidizing interface as $\exp(-X/L_D)$, where L_D is the decay length.

Chlorine additions to the oxidation ambient will tend to reduce the excess interstitials flowing into the substrate. Murarka [1.86] proposes that chlorine reacts with silicon at the interface to produce vacancies that in turn may recombine with oxidation-generated interstitials,

$$Si + \frac{u}{2} Cl_2 \overset{\rightarrow}{\leftarrow} SiCl_u + Si_V \tag{1.82a}$$

$$Si_V + Si_I \overset{\rightarrow}{\leftarrow} Si_S \tag{1.82b}$$
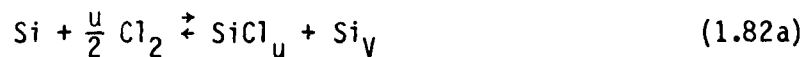
Fair [1.75] contends that u = 1 in reaction (1.82a) and therefore,

$$C_I \propto (p_{Cl_2})^{-0.5} \propto (p_{HCl})^{-0.5} \tag{1.83}$$

where p_{Cl_2} or p_{HCl} are the partial pressures of the chlorine addition to the ambient.

56

With the above considerations, the interstitial supersaturation with the oxidizing interface may be expressed as

$$\frac{C_I}{C_I{}^o} = \left(\frac{C_V{}^i}{C_V}\right) \left\{1 + \frac{K_4(T) \exp\left(-X/L_D\right)}{\left[\rho\, 1 + K_5(T)\, p_{HCl}^{0.5}\right]}\right\} \left(\frac{dX_o}{dt}\right)^{0.5} \tag{1.84}$$

where $K_4(T)$ and $K_5(T)$ are complex equilibrium and proportionality constants.

1.3.F.3.   Oxidation-Enhanced Diffusion

By combining eqns (1.75) and (1.84), the effective diffusivity under oxidizing conditions becomes

$$D = D_N + \Delta D_{ox} \tag{1.85a}$$

$$D_N = D_{SI} + f_{II}\, D_{SI}^i \left(\frac{C_V^i}{C_V}\right) \tag{1.85b}$$

$$\Delta D_{ox} = f_{II}\, D_{SI}^i \left(\frac{C_V^i}{C_V}\right) + \frac{K_4(T) \exp\left(-X/L_D\right)}{\rho\left[1 + K_5(T)\, p_{HCl}^{0.5}\right]}\left(\frac{dX_o}{dt}\right)^{0.5} \tag{1.85c}$$

The diffusivity has readily separated into two components. The first, D_N, is the impurity diffusivity under non-oxidizing conditions or a non-oxidizing surface. It contains both the vacancy-assisted term D_{SI} and the "equilibrium" non-oxidizing fractional interstitial-assisted contribution. D_N has in fact been modeled very well as an entirely vacancy-assisted process, as already discussed above for the commonly used dopants. The interstitial component has not been specifically considered, but apparently no serious failures have been discovered in most applications of a purely vacancy-based model under non-oxidizing surfaces. The interstitial component perhaps

57

makes only relatively minor contributions in such cases. With heavy doping, D_{SI} will increase substantially and the interstitial component apparently becomes even less significant relative to the increase of the total D. Thus, under non-oxidizing conditions--and especially with higher doping levels, D_N is satisfactorily modeled using eqns. (1.40), (1.42), (1.52), and (1.67) for B, As, P, and Sb, respectively

$$D_N \approx D_N \text{ (B, As, P, or Sb)} \tag{1.86a}$$

Under oxidizing conditions, however, interstitial generation adds an additional interstitial component ΔD_{ox} to the total diffusivity. Depending upon the oxidation conditions, this OED enhancement contribution may become comparable to and even exceed the D_N component to enhance the overall diffusivity D.

Note, however, that estimation of ΔD_{ox} depends upon the value for D_{SI}^i, the vacancy-assisted diffusivity under intrinsic non-oxidizing conditions. As already noted, models for D_{SI} have to this point neglected to distinguish a smaller interstitial component from the assumed dominant vacancy contribution, grouping both together as purely vacancy-assisted diffusion. Fortunately, if the interstitial component is comparatively small under intrinsic non-oxidizing conditions, then only very slight error will be introduced by assuming

$$f_{II} \ll 1 \tag{1.86b}$$

$$D_{SI}^i \approx D_N^i \tag{1.86c}$$

in calculating ΔD_{ox}. (This assumption will be shown to be consistent with experimental observations below.)

a. Oxidation Rate

OED of B and As has been observed to be greater in H_2O than in dry O_2 oxidation. Thus effect correlates with the higher oxidation rate

in H_2O that generates a greater interstitial supersaturation. ΔD_{ox} has been observed to approximate a $(dx_0/dt)^n$ power dependence, but the value of n has not been completely resolved. Taniguchi, et al. [1.81] found empirically that n = 0.3 for B and P in both dry O_2 and H_2O oxidations. Lin, et. al. [1.87] observed n ≈ 0.4 - 0.6 for B and P in dry O_2. This sublinear dependence is consistent with observations for (OISF) stacking fault growth [1.88], which has also been modeled to result from interstitial generation at the oxidizing interface [1.86]. From theoretical consideration of the possible reactions producing and consuming interstitials, Fair [1.75] and Hu [1.89] both conclude that n = 0.5. Thus, the assumption of n = 0.5 in the derivation of Eqn. (1.85c) for ΔD_{ox} appears reasonable. (However, recently Ishikawa, et al. [1.90] reported significantly higher values for n that also varied with the dopant: n ≈ 0.7-0.8 for phosphorus and n ≈ 0.9-1.1 for arsenic. This issue should be considered unresolved at present.)

b. Effect of Time-Averaging

The experimental observations of OED have been, of necessity, time-averaged over the total oxidation period to find an effective average enhancement $\Delta \tilde{D}_{ox}$. Given a power law $(dx_0/dt)^n$ dependence of the true instantaneous) enhancement ΔD_{ox}, a correction may be used. The experimental observations were generally for relatively long oxidation times and thick oxides, for which the linear-parabolic oxidation behavior is reasonably approximated as parabolic growth.

$$\frac{dx_0}{dt} \approx \left(\frac{B}{2X_0}\right) = \frac{1}{2}\left(\frac{B}{t}\right)^{0.5} \tag{1.87}$$

$$\Delta \tilde{D}_{ox} = \frac{1}{t_0}\int_0^{t_0} \Delta D_{ox}\ dt = \left(\frac{2}{2-n}\right)\Delta D_{ox} \tag{1.88}$$

59

Thus, with n = 0.5,

$$\Delta D_{ox} = \left(\frac{2-n}{2}\right) \Delta \tilde{D}_{ox} = \frac{3}{4} \Delta \tilde{D}_{ox} \qquad (1.89)$$

c. Orientation and Proportionality Constant

Surface orientation will determine the kink density ρ that in turn will reduce the interstitial concentration through surface regrowth. Fair [1.75] estimates $\rho \approx 10^{12}$ cm^{-2} for $\langle 100 \rangle$ silicon. Thus, by comparison of a similar derivation for OISF growth, based on the interstitial supersaturation, with stacking fault data for 100% dry O_2 oxidations [1.75], and including the time-averaging correction from eqn. (1.89), the proportionality factor $K_4(T)$ in ΔD_{ox} is estimated

$$\frac{K_4(T)}{\rho(100)} = \frac{(1.5)K_4(T)}{\rho(111)} = 1.31 \times 10^{-5} \; \exp\left(\frac{+2.82eV}{kT}\right) \; \frac{cm}{sec}^{-0.5} \qquad (1.90)$$

Thus, OED enhancement is greater for $\langle 100 \rangle$ than for $\langle 111 \rangle$ silicon, as experiments observed [1.91]. Also, as seen via the negative effective "activation energy" of the proportionality factor, the enhancement is greater at lower temperatures, also as demonstrated experimentally [1.80,1.81].

d. Fractional Interstitial Contribution and Impurity Dependence

The fraction of the substitutional flux that occurs by an interstitial-assisted mechanism is f_{II} for lightly doped silicon in non-oxidizing ambients. From the formulation of Eq. (1.85), f_{II} also is a direct measure of the diffusivity enhancement factor produced by the oxidation-generated interstitial supersaturation. This factor will vary with the impurity under consideration. Fair [1.75] has suggested that f_{II} in fact should correlate with the energy required to create an interstitial impurity atom. The efficacy of the interstitial-assisted mechanisms should be directly related to

the ease with which the impurity may be displaced into an interstitial site (by a silicon interstitial). Mizus and Higuchi [1.82] have noted that impurities with smaller ionic radii should therefore show a greater tendency to diffuse via interstitials. Thus, because the ionic radii vary as

$$B < P < As \approx Si \tag{1.91}$$

one should expect f_{II} to vary in inverse order.

Indeed, Fair's anaylsis [1.75] has produced the quantitative ratios at oxidation temperatures (~1000-1100°C) for f_{II}:

$$f_{II} (B : P : As) = 0.17 : 0.12 : 0.09 \tag{1.92}$$

However, note that f_{II} may not remain constant with temperature. It is a parameter indicating the relative magnitudes of interstitial vs. vacancy-assisted diffusion mechanisms. These two processes will likely have different activation energies. Thus, f_{II} as the ratio will also show a temperature dependence. Further, the diffusion energies vary with the impurity. Again, the f_{II} temperature dependence may then quite possibly also vary with the impurity.

The effective diffusivity under oxidizing conditions in eqn. (1.85) has been fitted to the OED data of Lin, et al. [1.80] in dry O_2 oxidation of <100> silicon to obtain values of f_{II} for boron, phosphorus, and arsenic.

$$f_{II}(B) = 4.09 \exp\left(\frac{-0.48 \ eV}{kT}\right) \tag{1.93a}$$

$$f_{II}(P) = 5.50 \exp\left(\frac{-0.57 \ eV}{kT}\right) \tag{1.93b}$$

$$f_{II}(As) = 6.85 \exp\left(\frac{-0.63 \ eV}{kT}\right) \tag{1.93c}$$

A slight temperature dependence is found that also varies somewhat with the impurity. The ratios of f_{II} for the impurities are in good agreement with

61

Fair's results [1.75] at comparable temperatures. The relative order of OED enhancement of the impurities is determined by the f_{II} values, decreasing in order from B to P to As as seen experimentally.

The values of f_{II} are comparatively small at lower temperatures, for which the OED enhancement is most significant. Thus for T $\leqslant 1000°C$,

$$f_{II} \leqslant 0.1 \qquad (1.94)$$

consistent with the assumption eqn. (1.86b,c) made in fitting the data. However, at higher temperatures, the intrinsic interstitial contribution becomes comparable to the vacancy-assisted component, and the f_{II} values become for T > 1200°C ～

$$f_{II} \leqslant 1.0 \qquad (1.95)$$

Assumptions in eqn. (1.86b,c) thus become very questionable. Fortunately, at these higher temperatures the OED enhancement factor also becomes less significant ($K_4(T)$ becomes negligible). Errors introduced by assumptions in eqn. (1.86b,c) in calculating ΔD_{ox} thus become unimportant.

e. Heavy Substrate Doping

With increasing substrate doping, as the silicon becomes extrinsic, (C_V/C_V^i) becomes greatly enhanced. The vacancy-assisted D_{SI} and therefore D_N increase correspondingly. However, the increased vacancy concentration should reduce the interstitial supersaturation by recombination. The OED enhancement ΔD_{ox} thus decreases as (C_V/C_V^i) even as D_N increases. These expectations have been confirmed experimentally, in good quantitative agreement [1.75] with the formulation for ΔD_{ox} in eqn. (1.85c), by the data of Raniguchi, et al. [1.81] for oxidation of silicon heavily doped with both B and P.

f.   Interstitial Decay Length

By examining the effect of increasing depth x from the inter-
face on the OED enhancement, Taniguchi, et al. [1.81] concluded that the
interstitial supersaturation extended considerable distances into the bulk.
The effective interstitial diffusion decay length L_D showed some slight tem-
perature dependence, decreasing with temperature from ~30 μ at 1000°C to
~25 μ at 1100°C. Thus, a reasonable, conservative assumption for the lower
temperature range at which ΔD_{ox} is significant is

$$L_D = 25 \ \mu \tag{1.96}$$

g.   Chlorine Oxidation

As discussed above, the OED enhancement should vary as
$(p_{HCl})^{-0.5}$ for chlorine additions to the ambient. Chlorine at the inter-
face is expected to generate vacancies that may then recombine with intersti-
tials to reduce C_I and ΔD_{ox} [1.75,1.86]. Limited data is available for
diffusion of boron and phosphorus in HCl/O_2 ambients from the work of
Nabeta, et al. [1.92]. They observed that chlorine additions at 1100°C and
1150°C did indeed reduce ΔD_{ox} to ~0 and D to ~D_N for HCl/O_2 in
excess of a few percent. However, at 1000°C increasing HCl/O_2 up to 10%
produced little change in ΔD_{ox}, and reduction of D to ~D_N did not occur.

Fitting eqn. (1.85) for D and ΔD_{ox} to this limited data yields reason-
able agreement with

$$K_5(T) = 2.31 \times 10^{34} \exp\left(\frac{-8.88 \ eV}{kT}\right) (atm)^{-0.5} \tag{1.97}$$

The activation energy of ~9 eV appears very high. It is apparent that the
influence of chlorine on OED is complex and inadequately understood, as is the
influence of chlorine on oxidation kinetics (see the discussion of chlorine

oxidation kinetics earlier in this report). Additional data for OED in chlorine oxidation is needed.

1.3.F.4. Silicon Vacancy Annihilation

The analysis above has neglected disturbances to the silicon vacancy population due to the oxidizing interface, tacitly assuming equilibrium vacancy concentrations. Undoubtedly, however, the oxidation process reduces vacancy levels both directly by vacancy consumption in the interface oxidation reaction and indirectly by recombination with the excess interstitials generated. A vacancy undersaturation relative to non-oxidizing conditions should result.

Unfortunately, understanding of the annihilation kinetics of vacancies during oxidation is even less than that of interstitial generation. As vacancy annihilation has been postulated as a step in the interface oxidation reaction [1.81, 1.7], the vacancy consumption rate may perhaps be directly related to the oxidation rate, possibly also by a power law similar to that proposed for interstitial generation. Other factors are known to influence vacancy levels as well. For example, heavy doping enhances vacancy concentrations; it also enhances diffusivity and oxidation rate, in direct correlation with those enhanced vacancy levels [1.67, 1.7, 1.23]. Recombination with generated excess interstitials may reduce vacancy concentrations. The disturbance of vacancy concentrations from equilibrium due to oxidation may be expected to decay with depth into the bulk, perhaps with a comparable decay length to that for the excess interstitials.

Clearly, the interaction of vacancies with interstitials and the oxidizing interface is complex. The direct dependence of vacancy concentration on oxidation rate is not known. The dependence on heavy doping and interstitial recombination may be approximated by

64

$$\frac{C_V}{C_V^{io}} = \left(\frac{C_V^O}{C_V^{io}}\right)\left(\frac{C_V}{C_V^O}\right) \ll \left(\frac{C_V^O}{C_V^{io}}\right)\left(\frac{C_I^O}{C_I}\right) \qquad (1.98)$$

where (C_V^O/C_V^{io}) represents the variation of vacancy concentration with heavy doping under non-oxidizing conditions, and (C_V/C_V^O) is the reduction in vacancies due to recombination with excess interstitials generated by the oxidation process. This mass-action driving force (C_I^O/C_I) has been used to approximate the (C_V/C_V^O) factor in analogy with eqn. (1.81). Again, the time constant required to achieve equality of these two quantities is not known. The extent to which this approximation is valid in the transient, quasi-steady-state situation of on-going oxidation is also not known. Clearly there is much work remaining to be done in this area.

1.3.F.5. Oxidation-Retarded Diffusion

Allowance for vacancy consumption during oxidation yeilds a modified form of the effective diffusivity in eqn. (1.75),

$$D \quad D_{SI}^i \left(\frac{C_V^O}{C_V^{io}}\right)\left(\frac{C_I^O}{C_I}\right) + \quad f_{II} \ D_{SI}^i \left(\frac{C_I}{C_I^O}\right) \qquad (1.99a)$$

$$D = D_{SI} \left(\frac{C_I^O}{C_I}\right) + \quad f_{II} \ D_{SI}^i \left(\frac{C_I}{C_I^O}\right) \qquad (1.99b)$$

For B, P, and As, oxidation produces OED, indicating that the enhancement of interstitial-assisted diffusion due to interstitial supersaturation more than compensates for reduction of vacancy-assisted diffusion due to vacancy undersaturation. This implies that f_{II} for these impurities must be sufficiently large. In fact, f_{II} must be significantly greater than the values in eqn. (1.93) that were derived assuming no vacancy consumption during

65

oxidation. Large values of f_{II} are consistent with the small ionic radii of these dopants relative to silicon.

Antimony, however, is substantially larger than silicon. This suggests that this impurity should show little tendency to diffuse via interstitials; i.e., f_{II} (Sb) should be relatively very small. In this instance, then diffusion is almost entirely via vacanies. Eqn. (1.99) then becomes

$$D(Sb) \simeq D_{SI}^i \left(\frac{C_V^0}{C_V^{i0}}\right)\left(\frac{C_I^0}{C_I}\right) = D_{SI}\left(\frac{C_I^0}{C_I}\right) \qquad (1.100)$$

Reduction of vacancy-assisted diffusion during oxidation will overwhelm enhancement of the negligible interstitial-assisted diffusion; overall diffusivity will in fact be _retarded_. These expectations of ORD for Sb have been confirmed experimentally in the limited data of Mizuo and Higuchi [1.82] for Sb diffusion in both H_2O and dry O_2 ambients. The degree of reduction of D(Sb) was observed to be substantial, with as much as 50% or more reduction in D(Sb), in reasonable agreement with eqn. (1.100).

It appears reasonable, then, to use eqn. (1.100) to estimate Sb diffusion under oxidizing conditions. Eqn. (1.84) for the interstitial supersaturation (C_I/C_I^0) may be used with slight modification. (C_V^i/C_V) in eqn. (1.84), representing the recombination driving force, has already been explicitly included in eqn. (1.100) as (C_V^0/C_V^{i0}) and should not be carried in substituting from eqn. (1.84) into eqn. (1.100). Still, understanding of the vacancy undersaturation during oxidation remains poor, and considerably more work is needed. When such models are developed, then the impurity diffusivities of eqn. (1.40), (1.42), (1.52) and (1.85) for B, P, and As, and eqn. (1.67) and (1.100) for Sb will demand major reworking.

1.4. EPITAXY

The growth of epitaxial layers is an essential part of bipolar processes and some MOS processes. It is, therefore, important for a general purpose process simulator to accurately model this technology. Such modeling is complicated by the dynamics of how growth rate and doping levels in epitaxial layers are related to gas phase silicon and dopant atom concentrations. SUP-REM Versions 1 and II avoid these questions by assuming a simple segregation model of the gas solid interface as described below.

The general situation is depicted in Fig. 1.21. Impurity redistribution during epitaxial growth includes both a diffusive component as well as a seg-regation effect. It is assumed that the bulk gas phase has a uniform equivalent concentration of single dopant species C_{gI}^*. This is a fictitious concentration equal to the desired uniform epi-doping concentration.

At the solid gas interface, there exist two flux components, F_S and F_b as indicated in Fig. 1.21. F_S is an impurity "evaporation" flux,

$$F_s = h(kC_{gI}^* - C_I) \qquad (1.101)$$

where h is the impurity evaporation coefficient. As evaporation flux, we define any impurity flux exchange between the gas and the solid, other than direct moving boundary incorporation; C_I is the impurity concentration at the solid surface, and k is the surface equilibrium segregation coefficient defined as

$$k = \frac{C_I}{C_{gI}^*} \qquad (1.102)$$

The term F_b is a flux induced by the interface motion and can be calculated as in the dase of thermal oxidation. It is given by
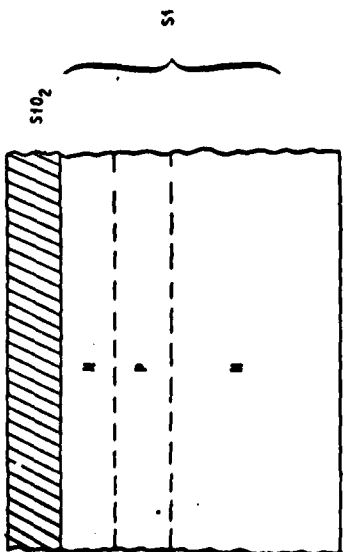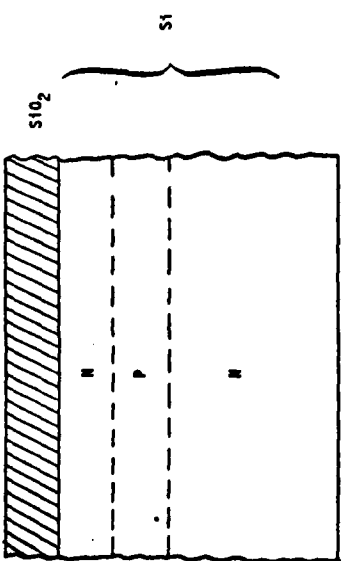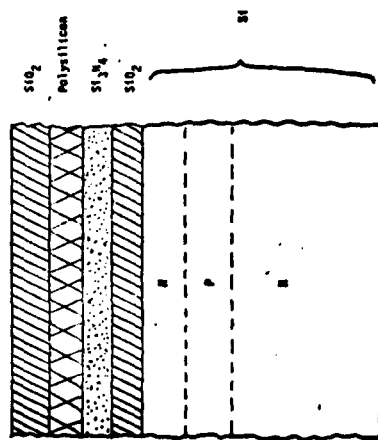
67

$$F_b = v(C_{gI}^* - C_I) \qquad (1.103)$$

where v is the interface velocity (growth rate). Within the solid (silicon) body, diffusive flow of impurities is accounted for as has been discussed previously.

A substantial amount of work, under this program over the past several years, has been aimed at improving the above model for epitaxial growth. A specific objective has been to understand the physical mechanisms responsible for transport of silicon and dopant atoms from the gas phase to appropriate lattice sites on the substrate. An accurate and general purpose (i.e. not system specific) physiochemical model which describes the growth rate and incorporation of dopant atoms into silicon epitaxial films during deposition from a $SiH_4-As_3-H_2$ mixture has been developed. Table 1.6 summarizes the older models in SUPREM versions 1 and II and the new model, described below, which is implemented in SUPREM III [1.93].

Epitaxial films with uniform impurity distribution are often required in IC fabrication. However, the distribution of impurities near the film-substrate interface is, in general, nonuniform. During the initial stages of growth the dopant incorporation process goes through a transient period, and 2-3 min are required [1.94] before the steady-state epitaxial doping level is established. As a result, a transition layer corresponding to this initial transient develops, and the epitaxial dopant concentration within this layer is either higher or lower than expected, depending on the initial conditions prior to the deposition cycle (e.g., substrate surface concentration, prebake time, prebake temperature, etc.). A special case of this initial transient problem is the commonly known "autodoping" phenomena, which occurs when lightly doped films are deposited on heavily doped buried-layers/substrates.

Table 1.6 - EPITAXY MODELS

SUPREM 1

2 layers
1 interface
2 materials

$F_{EVAP} = h(kC_G - C_I)$

$F_{GROWTH} = v(C_G - C_I)$

(Steady State Assumed)

SUPREM II

2 layers
1 interface
2 materials

$F_{EVAP} = h(kC_G - C_I)$

$F_{GROWTH} = v(C_G - C_I)$

(Steady State Assumed)

SUPREM III

5 layers
4 interfaces
Any materials

Numerical solution of Fick's Second Law; boundary conditions model transient growth kinetics

$$\frac{\partial N(x,t)}{\partial t} = D \frac{\partial^2 N(x,t)}{\partial x^2}$$

$$D \left. \frac{\partial N}{\partial x} \right|_{x = \infty} = 0$$

$$D \left. \frac{\partial N}{\partial x} \right|_{x = surface} = f(t)$$

where f(t) involves dopant and silicon partial pressures and reaction physics.

69

The extent of this transition layer imposes severe limitations on the fabrication of submicron epitaxial films, which will be required with the development of VLSI technology.

Epitaxial films with intentional, non-uniform impurity distributions are sometimes also required. These can be used to fabricate specialized devices or to optimize device characteristics. Examples of these are the hyper-abrupt epitaxial tuning diodes, in which precisely controlled impurity distributions are needed in order to obtain the specific C-V characteristics. Non-uniform doping profiles can be fabricated by varying with time the dopant gas flow entering the epitaxial reactor during deposition.

The general problem is that of determining a computer model capable of simulating the epitaxial doping profile resulting from any given time-varying, gas-phase composition during growth. The model must also account for thermal redistribution of dopant atoms in silicon during epitaxial deposition. Because the redistribution of impurities within the solid is controlled by diffusion, the starting mathematical framework is to solve Fick's Second Law throughout the solid silicon, from a plane very deep inside the substrate up to the silicon surface. That is (see Fig. 1.22)

$$\frac{N(z,t)}{t} = D \frac{\delta N(z,t)}{\delta z^2} \qquad > z > z_f \qquad (1.104)$$

where N is the dopant concentration in the solid, D is the diffusion coefficient of the impurity in solid silicon, and z and t are the spatial and time variables, respectively. As shown in Fig. 1.22, the z-direction goes perpendicular to the silicon surface, and z_f is defined as the location of the moving gas-solid interface.

The solution to the diffusion equation during dpitaxial growth must satisfy the following initial and boundary conditions. The initial condition is given by

$$N(z,0) = f_1(z) \qquad (1.105)$$

where $f_1(z)$ represents the distribution of impurities in the substrate just before epitaxial deposition.

The boundary conditions are

$$D \left. \frac{\delta N}{\delta z} \right|_{z \to \infty} = 0 \qquad (1.106)$$

$$D \left. \frac{\delta N}{\delta z} \right|_{z=z_f} = f^2(t) \qquad (1.107)$$

Equation (1.106) indicates that the impurity diffusive flux at a plane very deep in the silicon substrate is zero. This is a reasonable boundary condition because silicon wafer thicknesses are much larger than diffusion lengths in bulk silicon.

Equation (1.107) indicates that the impurity diffusive flux in the solid at the gas-solid interface is, during epitaxial growth, a function of time. As will be shown later, an expression for $f^2(t)$ can be derived from a basic understanding of the mechanisms controlling the incorporation of impurities in the epitaxial silicon during growth. The time-dependence of Eq. (1.107) is related to (i) the transients associated with the establishement of a steady-state deposition process, and (ii) the time-variation (if any) of the gas-phase composition in the reactor. Notice that the autodoping problem falls in the first category. The right-hand side of Eq. (1.107) is a function of dopant partial pressure, epitaxial growth rate, deposition temperature, and to a lesser degree: reactor geometry, hydrogen velocity, etc. In order to solve

Fick's Second Law during epitaxial growth an expression for $f_2(t)$ in Eq. (1.107) must be obtained. This is accomplished by using the epitaxial doping model described below.

In this section, the behavior of dopant species in each important region of an epitaxial reactor is studied in detail, both in terms of gas-flow dynamics and chemical kinetics. Figure 1.23 shows schematically a section of the horizontal reactor tube along the length of the susceptor. Three main regions are indicated in Fig. 1.23: 1) main gas stream, 2) boundary layer, and 3) adsorbed layer. The main gas stream consists of hydrogen mixed with silane and arsine flowing by forced convection. The boundary layer is a thin layer, about 5 mm thick, through which reactants diffuse to the silicon surface. The adsorbed layer consists of a population of hydrogen, silicon-, and dopant-containing species that occupy adsorption sites and are capable of moving on the solid surface. The behavior of dopant species in each of these three regions is now discussed, and equations describing this behavior will be derived. The equations will then by used to obtain an expression for $f_2(t)$ in Eq. (1.107).

1.4.A. Main gas stream.

In most epitaxial reactors, gas-phase depletion of dopant species in the main gas stream is almost negligible. Therefore, the partial pressure of dopant species in the main gas stream region can be assumed to be independent of position, and nearly equal to the partial pressure of dopant species at the reactor input. Moreover, the time constant associated with the transport of dopant species in the main gas stream by forced convection is of the order of 1 sec [1.94], while the time constant of the overall doping process in a horizontal reactor is of the order of 40 sec [1.95]. Therefore, it can be assumed that any time-variation of the dopant partial pressure at the reactor input is transmitted "instantaneously" throughout the main gas stream.

72

Consequently, the following expression can be used to describe the behavior of dopant species in the main gas stream region

$$P_D(t) \quad P_D{}^\circ(t) \tag{1.108}$$

in which P_D is the partial pressure of dopant species in the main gas stream region, and $P_D{}^\circ$ is the input dopant partial pressure.

1.4.B.   Boundary layer

The time constant associated with the transport of dopant species through the boundary layer by diffusion is of the order of 0.1 sec [1.95], which is much shorter than that of the overall doping process. Therefore, it is reasonable to assume that the flux of dopant species leaving the boundary layer by adsorbing on the silicon surface (F_S) follows closely any time-variation of the dopant flux entering the boundary layer from the main gas stream.

$$F_S(t) \quad F_z(\bar{z},t) \tag{1.109}$$

The two fluxes in Eq. (1.109) can be expressed in terms of deposition parameters. The flux of dopant species leaving the main gas stream toward the wafer surface $[F_z(\bar{z})]$ can be approximated by [1.96]

$$F_z(\bar{z}) = k_m[P_D{}^\circ - P_D{}^*] \tag{1.110}$$

in which k_m is the boundary layer mass transport coefficient of dopant species in hydrogen, and $P_D{}^*$ is the dopant partial pressure just above the gas-solid interface. k_m is a function of reactor geometry, hydrogen velocity, gas-phase temperature, etc. An expression for $F_S(t)$ is obtained in the next section.

1.4.C.   Adsorbed layer

In order to determine F_S, the sequence of steps taking part in the doping process and occurring at the silicon surface must be considered in detail. These steps are shown in Fig. 1.24 [1.96]. When an arsine molecule

73

in the gas phase is close to the silicon surface it undergoes a process of adsorption [step (1)]. The arsine molecule, once adsorbed, decomposes chemically yielding elemental arsenic [step (2)] which then diffuses on the surface until it finds an incorporation site and attaches to it [step (3)]. The arsenic atom, now incorporated in the silicon lattice, is quickly covered by subsequently arriving silicon atoms [step (4)], and diffuses into (or out of) the bulk silicon. For the treatment presented here, it is sufficient to select one of the above steps as the rate-limiting step, leaving all other steps in the sequence near thermodynamic equilibrium [1.95]. If step (1) above is assumed to be the slowest in the sequence, the surface steps can be summarized in the following two reactions

(i) Arsine adsorption

$$AsH_3(g) + s \rightleftharpoons AsH_3 - s; \quad K_1 = \frac{(\theta_D N_s)}{P_D^* (\theta N_s)} = \frac{k_F}{k_R} \tag{1.111}$$
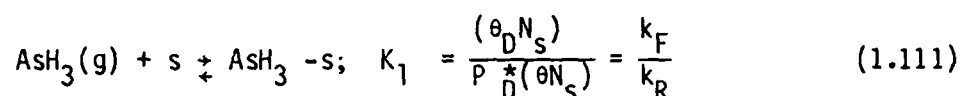
(ii) Chemical decomposition, site incorporation, and covering of arsenic by silicon

$$AsH_3 - s \rightleftharpoons As(ss) + s + \frac{3}{2} H_2(g)$$

$$K_2 = \frac{(k_H N)(\theta N_s) P_{H_2}^{3/2}}{(\theta_D N_s)} \tag{1.112}$$
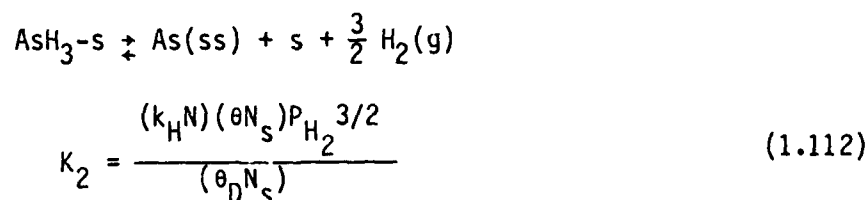
where s represents a vacant adsorption site on the surface, AsH_3-s represents an arsine molecule occupying an adsorption site, K_1 and K_2 are equilibrium constants, θ_D is the fraction of adsorption sites occupied by arsine, N_s is the surface density of adsorption sites per unit area, k_F and k_R are the forward and reverse reaction rate constants for arsine adsorption, θ is the fraction of adsorption sites which are vacant, k_H is

74

Henry's law constant, and P_{H_2} is the hydrogen pressure (~ 1 atm). An expression for F_S can now be obtained by using Eq. (1.111) and (1.112)

$$F^S = k_f[P_D^* - N/K_p] \qquad (1.113)$$

where $k_f \equiv k_f(\theta N_s)$ is a kinetic constant associated with arsine adsorption and $K_p = K_1 K_2/k_H$ is a thermodynamic constant relating dopant species in the gas phase and the solid silicon. Equations (1.110) and (1.113) can now be used to obtain the following expression relating F_S and the input dopant partial pressure $P_D°$

$$F^S = k_{mf}[P^{D°} - N/K^P] \qquad (1.114)$$

in which $k_{mf} \equiv [1/k_m + 1/k_f]^{-1}$.

As discussed earlier, the times associated with gas-phase mass transport of dopant species are negligible as compared to the time constants measured for the overall doping process. Therefore, the only mechanisms that could be responsible for these long time constants ought to be associated with the adsorbed layer.

By considering mass balance of dopant species in the adsorbed layer the following equation is obtained

$$F_S - gN(z_f) + D \left.\frac{\delta N}{\delta z}\right|_{z=zf} = \frac{d(\theta_D N_s)}{dt} \qquad (1.115)$$

where g is the epitaxial growth rate. In Eq. (1.115), F_S represents the rate at which the adsorbed layer increases its population of dopant species. The second term in Eq. (1.115) represents the rate at which the adsorbed layer decreases its population of dopant species due to the silicon covering step (see Fig. 1.24), and the third term represents diffusive exchanges between dopant atoms in the adsorbed layer and the bulk silicon. The right-hand side of Eq. (1.115) represents the rate of change of the density of dopant species

75

per unit area in the adsorbed layer, and becomes zero when the overall doping process reaches steady state. By substituting Eq. (1.114) into Eq. (1.115), the following expression is obtained

$$k_{mf}[P_D^\circ - N(z_f)/K_p] - gN(z_f) + D\frac{\delta N}{\delta z}\bigg|_{z=zf} = K_A \frac{\delta N(z_f)}{\delta t} \qquad (1.116)$$

in which Eq. 1.112 was used with $K_A \equiv k_H(\theta N_s)P_{H_2}3/2/K_2$. Eqn. 1.116 relates the epitaxial dopant concentration at the silicon surface $N(z_f)$ and the input dopant partial pressure P_D°. It is clear from Eq. (1.115) and (1.116) that any abrupt variation of P_D° with time during epitaxial growth is not transmitted "instantaneously" to $N(z_f)$ because some time is required before the population of dopant species in the adsorbed layer $\theta_D N_s$ accommodates to the new steady-state condition. This storage-like behavior of the adsorbed layer is responsible for the relatively long time constants measured experimentally for the overall doping process. Equation (1.116) can be rearranged to give the boundary condition $f_2(t)$ needed in Eq. (1.107), i.e.

$$D\frac{\delta N}{\delta z}\bigg|_{z=zf} = f_2(t) = -k_{mf}[P_D^\circ - N(z_f)/K_p] + gN(z_f) + K_A\frac{dN(z_f)}{dt} \qquad (1.117)$$

Fick's Second Law (Eq. 1.104) can now be solved subject to the initial and boundary conditions given by Eq. (1.105) and (1.106) and (1.117), respectively. Values for the parameters in the doping model (k_{mf}, K_p, K_A) can be determined as described in [1.93].

Figures 1.25a and b show examples of the simulation capability of this new model. The figures show both measured and simulated doping profiles corresponding to the time-varying arsine flows indicated in the insets. In both cases the input to the epitaxial system consisted of an increaasing step in arsine flow followed by a decreasing step, the combination simulating a pulse.

76

The epitaxial layer in Fig. 1.25a was deposited at 0.33 μm/min, and the pulse duration was 8.6 min. The epitaxial layer in Fig. 1.25b was deposited at 0.59 μm/min, and the pulse duration was only 48 sec. The broken line shows the impurity profile simulated by SUPREM III. The solid line shows the corresponding majority-carrier distribution obtained by using the SUPREM III output and calculating a one-dimensional solution to the Poisson's equation. The computer program SEDAN was used to obtain the Poisson solution [1.97]. The circles indicate the epitaxial doping profile as determined by capacitance voltage measurements on deep-depletion MOS structures [1.95]. The agreement between the simulated majority-carrier profile and the measured profile is better than that between the SUPREM III output and the measured profile. This is because the profile measured by the capacitance-voltage technique is a closer representation of the majority-carrier distribution than it is of the impurity distribution [1.98]. The experiments shown in Fig 1.25 were carried out in a horizontal reactor (Hugle Model HIER II). The parameters used to simulate the Hugle reactor were obtained from the data in [1.95], and are shown in Table VII together with those corresponding to the Unipak unit used in other studies reported in earlier annual reports. Table 1.8 shows values of decay time and decay length L [1.94,1.95] for each reactor at different silicon growth rates, calculated from

$$\tau \cong \frac{K_A}{g + \frac{k_{mf}}{K_p}} \qquad (1.118)$$

$$L \cong g\tau \qquad (1.119)$$

It is interesting to observe that τ and L of each reactor (for a given growth rate) are remarkably close, which indicates that the extent of any transient phenomena (including autodoping) is independent of the horizontal reactor

77

used. These experiments clearly indicate that SUPREM is capable of simulating epitaxial doping profiles corresponding to different time-varying dopant gas flows. Moreover, these experiments demonstrated that the simulation capability of SUPREM is not restricted to one single reactor.

	Unipak	Hugle
K_p (cm^{-3} atm^{-1})	1.05×10^{26}	3.8×10^{25}
k_{mf} ($cm^{-2}sec^{-1}$ atm^{-1})	4.85×10^{19}	5.26×10^{18}
K_A (cm)	5.7×10^{-5}	4.1×10^{-5}

Table 1.7. Reactor parameters corresponding to the Unipak and the Hugle units

	Unipak		Hugle	
	τ (sec)	L (μm)	τ (sec)	L (μm)
0.3 μm/min	59	0.3	64	0.32
0.45 μm/min	47	0.35	46	0.35
0.6 μm/min	39	0.39	36	0.36

Table 1.8. Decay time τ and decay length L calculated for the Unipak and the Hugle reactors at different silicon growth rates

1.5. POLYSILICON

CVD deposition of polysilicon is of major importance in modern MOS technologies and is finding increased use in bipolar structures as well. In addition to film thickness and dopant masking properties, thermal oxidation rates and electrical resistivity are of prime importance in correctly simulating such structures. A large amount of work, primarily concentrating on resistivity, has been undertaken under this program in the past few years. We have included, for the first time in SUPREM III, a modeling capability for polysilicon.

From a user's point of view, the program should ideally function as illustrated in the flow chart below.

User Input:

Deposit 5000 A poly, 600°C, LPCVD

↓

Program Calculates:

Grain Size, resistivity

↓

User Input:

a) Oxidize - T, t, ambient
b) Dope - Dose, energy or predep
c) Anneal - T, t, ambient
d) Etch - t, conditions

↓

Program Calculates:

Grain size, resistivity, oxide thickness

In order for SUPREM to accomplish this, accurate process models are required for the following steps.

(1) Grain growth as f (T, t, doping level)

(2) Oxidation kinetics as f (T, t, ambient, doping level)

(3) Dopant diffusion in poly as f (T, t, doping level)

79

(4) Dopant segregation at grain boundaries as f (T, t, doping level)

(5) Carrier trapping at grain boundaries as f (T, t, doping level)

(6) Carrier mobility

At the present time, basic models exist for each of these steps. Some of these models are more developed than others, as will become apparent in the paragraphs below. Further work is needed to refine the oxidation and diffusion models, in particular, however. The basic models which are implemented, should provide a reasonable simulation capability for SUPREM III. It is our intent to continue work on these models, and to provide updates for SUPREM III as improvements become available. In the paragraphs below, a summary of the models as they now stand is provided.

1.5.A.   Grain Growth

The grain-growth mechanism in poly Si is believed to be controlled by diffusion. This movement can be accomplished by a series of diffusion jumps of individual atoms across the grain boundaries, and the growth rate is then determined by the diffusion rate of the atoms across each grain boundary. In the absence of a driving force, the number of jumps in the opposite direction is equal and, as a consequence, there is no net movement of atoms. When a driving force exists, the free energy is reduced and atoms migrate from one grain to another which contributes to the migration of the boundary and the resultant grain growth.

The driving force has the following two components: (1) energy difference between atoms at one side of grain boundary and at the other, and (2) interfacial energy between two grains. Because of this interfacial energy, the grain boundary tends to minimize its area, which makes the driving force inversely proportional to the boundary curvature. The driving force can be expressed by [1.99],

80

$$F = \frac{a\lambda b^2}{L} \qquad (1.120)$$

where λ is a grain-boundary energy, b is the lattice constant, L is the average grain size, and a is a geometric factor. The migration rate of the boundary is not only proportional to the external driving force but also to grain-boundary mobility. Thus mobility measures atom migration under an external driving force and is directly proportional to the Si self-diffusivity in a diffusion-controlled growth regime [1.99]. This leads to greater grain-boundary mobility at higher vacancy concentrations. The mobility is given by:

$$\mu = \frac{D^g}{kT} \qquad (1.121)$$

where D^g is the Si self-diffusion constant along a boundary.

The boundary migration rate is the product of mobility and the driving force,

$$\frac{dL}{dt} = \mu F = \frac{a\lambda b^2 D^g}{LkT} \qquad (1.122)$$

which when rearranged and integrated, results in:

$$L^2 = L_0^2 + 2 \int_0^t a\lambda b^2 \frac{D^g}{kT} \, dt \qquad (1.123)$$

If the grain-boundary energy and Si self-diffusivity do not vary with time, the final grain size can be expressed in terms of the initial grain size L_0 and a grain-growth rate constant k_r,

$$L = \sqrt{L_0^2 + k_r^2 \, t} \qquad (1.124)$$

where k_r is

$$k_r = \frac{2a\lambda b^2 D^g}{kT}$$

When the initial grain size is small compared to $\sqrt{k_r^2 t}$, this reduces to $L \approx k_r \sqrt{t}$ which is valid when the poly-Si is deposited at temperatures much lower than the annealing temperature. Because the active doping concentration changes with the grain size and out-diffusion, the following equation is used in the ith time step to determine the grain growth with the new rate constant.

$$L_{i+1} = L_i + \left(\frac{ab^2}{kT}\right) \frac{\lambda_i D_i^g}{L_i} (t_{i_{ti}} - t_i) \qquad (1.125)$$

where the subscript i indicates the ith time step. The model to calculate the change of active doping concentration, which determines D_i^g as a function of grain size due to segregation at the grain boundaries will be discussed in a later section. The grain-boundary free energy λ_i, which changes as a function of the polysilicon layer thickness and the amount of impurity segregation at grain boundaries, is determined empirically.

Because of the energy balance, grain boundaries are the preferred sites for mobile defects, primarily vacancies. The silicon self-diffusion constant at grain boundaries can be related to the bulk self-diffusivity by [1.100]:

$$D^g = D^l \left(\frac{D_{in}^g}{D_{in}^l}\right) \Big/ \left[1 + AC_1 \exp \frac{G_a}{kT}\right]^2 \qquad (1.126)$$

where G_a is heat of segregation, A is the vibrational entropy factor, D^l is the diffusivity in bulk, here the subscript in denotes the intrinsic polysilicon, and C_1 is the impurity concentration in an atomic fraction inside the grain determined by the segregation of impurities at the boundary, which is a function of grain size. The ratio of the two diffusivities can also be expressed by,

82

$$\frac{D_i^g}{D_i^l} = \frac{D_{io}^g}{D_{io}^l} \exp\left[\frac{\Delta G_b}{kT}\right]$$

$$= \gamma \exp\left[\frac{\Delta G_b}{kT}\right] \tag{1.127}$$

where ΔG_b is the difference in the activation energies between the grain
boundary and the bulk, γ is the ratio of pre-exponential factors of the two
diffusivities. The activation energy for grain-boundary diffusion is always
lower than that for diffusion through grains, which implies that ΔG_b should
be positive [1.101]. One significant consequence is that the relative impor-
tance of grain-boundary diffusion becomes less at higher temperatures. The
grain-boundary diffusivity D^g is given by,

$$D^g = D^l \gamma \frac{\exp\frac{\Delta G_b}{kT}}{\left[1 + AC_1 \exp\left(\frac{G_a}{kT}\right)\right]^2} \tag{1.128}$$

Here, bulk self-diffusivity D^l as a function of doping concentration can be
determined from the vacancy concentrations [1.101].

When the average grain size in polysilicon becomes comparable to the
thickness of the polysilicon layer, many grains are in contact with the oxide
layer and/or the Si substrate interface and the interfacial energy becomes
more important in the growth mechanism. Grains with an orientation having
lower interfacial energy will grow at the expense of grains with high inter-
facial energy. This will result in preferential orientation where some grains
can grow several times larger than others and this is believed to be the cause
of secondary recrystallization in polysilicon layers. This secondary recrys-
tallization has a quite different origin from that in the metal systems.

83

In SUPREM III, the thickness effect is described by modifying the grain boundary energy term according to an empirical equation,

$$\lambda = \frac{\lambda_0}{1 + h \frac{\text{interface area}}{\text{grain boundary area}}} \qquad (1.129)$$

where the constant h is determined from experimental res_lts. The interface area is always 2 per square of polysilicon, and the total grain boundary area per square of polysilicon is calculated by assuming that all grains are spherical and have the same size.

An example of the application of this grain growth model is shown in Fig. 1.26. Simulated and experimental results are shown for three different polysilicon film thicknesses, annealed at 950°C. Despite the effect of a variable active concentration throughout annealing, the increment of grain size follows a $\sqrt{t}$ dependence. The variations are not significant. The empirical parameter h in Eqn. (1.129) was determined to be approximately 3 by matching the simulated and experimental data.

1.5.B. Polysilicon Oxidation

Polysilicon films are widely used in MOS technologies today and are finding increasing use in bipolar structures as well. These films are generally deposited by CVD or LPCVD and are commonly subjected to thermal oxidation processes during device fabrication. In some applications (load resistors in static RAMS for example), the polysilicon films are lightly doped and exhibit oxidation kinetics intermediate between <111> and <100> single cyrstal kinetics. This behavior is to be expected because of the variety of grain orientations present in polysilicon films. Modeling of the oxidation kinetics in this case is relatively straightforward since the <111> and <100> orientations represent upper and lower bounds for oxidation rate in single

84

crystal substrates. In fact, modeling can proceed via the familiar linear

parabolic growth law

$$\frac{\left(X^2 - X_i{}^2\right)}{B} + \frac{\left(X - X_i\right)}{B/A} = t \qquad (1.130)$$

where

$$\frac{B}{A} \langle 111 \rangle \;>\; \frac{B}{A} \;\text{Poly}\; >\; \frac{B}{A} \langle 100 \rangle$$

and

$$B_{POLY} = B_{SINGLE\ CRYSTAL}$$

The parabolic rate constants are expected theoretically to be equal in

the two cases because the physical process represented by B is oxidant diffu-

sion through an amorphous SiO_2 layer. Thus whether the SiO_2 is grown on

single crystal (of any orientation) or on polysilicon, B is expected to be the

same. This has been confirmed experimentally. Since the B/A values for $\langle 111 \rangle$

and $\langle 100 \rangle$ single crystal are different by only a factor of $\simeq 1.7$, an average

value can be used for polysilicon with reasonably accurate results.

The major difficulty in modeling polysilicon oxidation kinetics occurs

when the films are heavily doped, as they often are when used for MOS gate

electrodes or for interconnects. The effect of doping on the oxidation kinet-

ics of single crystal silicon has also been investigated and accurate models

are available for this process. The most widely accepted model is due to Ho

and Plummer [1.7,1.8], according to which, the presence of dopant in the silicon

shifts the Fermi level, causing an increase in the number of vacancies, which

enhances the surface reaction between silicon and the oxidant. In addition, the

presence of dopant in the SiO_2 increases the diffusivity of the oxidant

leading to an increased oxidation rate through a change in the parabolic rate

constant B. This model works well for single crystal silion; however, it cannot

be expected to completely explain the oxidation behavior of polysilicon - mainly

because of the presence of grain boundaries. It is widely known that dopant

segregation [1.102], carrier trapping [1.103] and carrier tunneling [1.104] at the grain boundaries can markedly change the electrical properties of polysilicon. Therefore it is expected that grain boundaries should also play a very important role in the oxidation behavior of polysilicon, since the electrically active doping concentration determines the total vacancy concentration and hence B/A in [1.7, 1.8].

Irene et. al. [1.105] have indeed observed by TEM studies that there are thickness undulations in the oxide because of different oxidation behavior at the grain boundaries. Kamins and MacKenna [1.106] have reported that at lower oxidation temperatures where the oxidation mechanism is predominantly surface reaction controlled, the rate of oxidation of polysilicon is characteristic of the random orientation of the grains as described above. The different oxidation rates across the surface cause surface roughening. At higher oxidation temperatures where the oxidation is predominantly diffusion controlled, a smoother polysilicon oxide surface is produced. Most other work has been aimed at determining the doping enhanced oxidation of polysilicon and investigating processing conditions where in an MOS process, for example, for the growth of a given thickness of gate oxide, the thickness of the oxide on the polysilicon could be maximized. Sunami [1.107] studied steam oxidation of phosphorus doped polysilicon in the temperature range of 700-850°C, Kamins [1.108] studied steam oxidation of phosphorus doped polysilicon at 850°C and Barnes et. al. [1.109] studied oxidation of phosphorus doped polysilicon in dry O_2 in the temperature range of 750-900°C. Baldi et.al. [1.110] investigated steam oxidation of arsenic doped polysilicon at 850°C, where doping was done during the growth. In most of these studies no attempt was made to understand the basic kinetics of the process so that a generalized model could be developed for the thermal oxidation of polysilicon.

Under this contract, a series of experiments have been underway for some time to try to resolve some of the questions regarding the kinetics of oxidation of heavily doped films [1.16]. Fig. 1.27 is representative of some of the results. Additional details are described in Section 4 of this report.

Several interesting observations can be made from this figure which indicate that additional mechanisms entering into the polysilicon oxidation process make it substantially different from that of single crystal silicon. For all oxidations the oxides grown on 0.5 μm thick polysilicon are consistently thicker than those grown on 1 μm thick films. The effect is stronger for thinner oxides.

It is obvious from the figure that heavy doping enhances the rate of oxidation of polysilicon relative to undoped single crystal silicon. The enhancement is stronger for lower oxidation temperatures. This behavior is similar to that observed for doped single crystal silicon; however, if the data in the figure are compared to that of the oxidation of phosphorus doped single crystal silicon [1.7, 1.8], it can be concluded that the enhancement of oxidation because of phosphorus doping is much greater for polysilicon than for single crystal silicon. In fact, it appears that a reasonable interpretation of the data for very heavily doped films is to let $B/A \rightarrow \infty$ and B = single crystal value.

After studying a significant body of data, we have concluded that a reasonable interim fit to the experimental results can be achieved in SUPREM III if we simply make use of the single crystal models for heavy doping kinetics, modified to account for the fact that not all the impurities in polysilicon are electrically active. Thus SUPREM III uses the grain growth model described earlier and the resistivity model to be described shortly, to calculate the electrically active dopant concentration in a polysilicon film and from

this n value, $(B/A)_{EFF}$ can be calculated from the models used for single crystal oxidation.

This approach is not rigorously correct because other phenomena apparently occur in polysilicon oxidation. However at the present time, this approach will allow reasonable accuracy in simulated oxide thicknesses We plan to update this approach in SUPREM III as new data and models become available.

1.5.C.   Polysilicon Resistivity

The model used to calculate the electrical resistivity of polysilicon for SUPREM III is described in this section. This model combines models for grain growth, dopant segregation and carrier trapping at the grain boundaries and therefore can predict the resistivity as a function of process parameters, such as temperature, time, doping density, etc. This resistivity model has resulted from a major effort over the past several years under this contract, involving both experimental work and theroetical analysis.

The resistivity of undoped and lightly doped polycrystalline-silicon is on the order of 10^5 to 10^6 Ω-cm which is several orders of magnitude higher than that of single-crystal silicon. Resistivity is fairly constant in this range, and changes in dopant concentrations result in only small variations in resistivity. At medium doping levels, however, resistivity is a strong function of dopant concentration and a slight increase will cause it to drop sharply. At high doping levels, resistivity approaches that of single-crystal silicon although it will always remain slightly higher.

Polycrystalline silicon is generally viewed as composed of small crystallites joined together by grain boundaries. Inside each crystallite, the atoms are arranged in a periodic manner forming small single crystals, while the

grain boundaries are composed of disordered atoms with incomplete bonding. The high concentration of defects and dangling bonds at the grain boundaries cause trapping states capable of immobilizing both dopant atoms and charge carriers, thus reducing the available carriers for conduction and their mobility. A conceptual view of the effects of carrier trapping and dopant segregation on the electrical conduction is shown in Fig. 1.28.

During high temperature processing, depending on dopant concentration, grain size, and temperature, some of the dopant atoms segregate to the grain boundaries where they are trapped and become electrically inactive. The remaining dopant atoms are distributed uniformly within the grains and can be ionized as in single-crystal silicon.

As appreciable fraction of the dopant atoms have been shown to segregate to grain boundares in arsenic- or phosphorus-doped polycrystalline silicon while no segregation has been observed in the case of boron [1.104]. The number of dopant atoms segregating to the grain boundaries has been shown to be a function of processing or annealing cycles, with more dopant atoms segregating to the grain boundaries upon annealing at lower temperatures. The mechanism of dopant segregation to the grain boundaries can be modeled by [1.104].

$$\ln \frac{N_{GB}}{N_G} = \ln \frac{AQ_s}{N_{SI}} + \frac{Q_0}{kT_A} \qquad (1.131)$$

where N_G and N_{GB} are the dopant concentration/cm³ in the grains and at the grain boundaries, respectively, Q_s is the density of grain boundary sites/cm³, N_{Si} is the number of Si atoms/cm³, Q_0 is the heat of segregation, $A = \exp(-S/k)$, S is the vibrational entropy, T_A is the annealing temperature and k is Boltzmann's constant.

After accounting for the fraction of the dopant atoms segregating the grain boundaries, the remaining dopant atoms are distributed uniformly within

89

the grains. The number of ionized dopant atoms in the grains can be deter-
mined using Fermi-Dirac statistics in a similar way as in single-crystal sili-
con. The details of this model have been described in [1.104].

The second phenomenon included in the model is carrier trapping at the
grain boundaries. The concept of carrier trapping is illustrated in Fig.
1.29. The theory of the carrier trapping phenomenon and its influence on
carrier conduction was initially developed by Seto [1.103] and subsequently
modified by Bacarani et. al. [1.111] and Lu et. al [1.112]. According to the
carrier trapping model, the high concentration of defects and dangling bonds
at the grain boundaries results in the formation of trapping states that are
capable of trapping and immobilizing the carriers, thereby reducing the number
of free carriers available for conduction. These traps capture the carriers
and become electrically charged. As a consequence of carrier trapping, a
portion of each grain is depleted, creating potential energy barriers, qV_b,
which impede the motion of carriers from one crystallite to another. Charge
transfer in this case is controlled by thermionic emission over these barriers
which have a maximum height given by

$$qV_b = q^2 Q_t^2 / 8 \ \epsilon N_G \qquad (1.132)$$

where Q_t is the density of trapping states/cm^2 at the grain boundary, ϵ is
permittivity of silicon and N_G is the density of dopant atoms from Eqn.
(1.131). In models described in [1.103, 1.111, 1.112] N_G has been assumed
to be the total average doping density. Segregation has been completely
neglected, resulting in an error in the case of P and As doped polysilicon. A
modified relationship based on [1.103, 1.111, 1.112] has been derived which
relates the thermionic emission current density J_{th} to the potential drop
across the grain boundary V_{GB}

$$J = \frac{q^2 N_g L}{kT} \left[\frac{kT}{2\pi m^*}\right]^{1/2} \exp\left(-\frac{qV_b}{kT}\right) V_{GB} \tag{1.133}$$

where N_g is the active carrier concentration in the grain, L is the grain size, k is Boltzmann's constant, m^* is the effective carrier mass and T is the measurement temperature. From this relationship the resistivity and the mobility of the grain boundary can be calculated. By assuming the properties of the undepleted grain to be the the same as that of single crystal silicon its resistivity can be calculated. By combining the resistivity of the grain and the grain boundary the average resistivity of the polysilicon film can be calculated.

An example of the application of this model to the calculation of poly resistivity is shown in Fig. 1.30. This particular example [1.113] considers Phosphorus doped poly annealed at 950°C for three hours. The poly layer was protected by a CVD oxide to prevent surface evaporation of impurities during the annealing. The fraction of the dose segregated at poly grain boundaries calculated by the model is shown along with calculated and measured resistivity. Reasonable agreement is apparent. Also shown in the figure is the average grain size, both calculated and measured by TEM. Again the agreement is good.

1.5.D.   Diffusion In Polysilicon/Silicon Structures

If SUPREM III is to accurately model the electrical properties and dopant profiles in polysilicon, it must model dopant diffusi   , within, and from these films. The general problem is illustrated in Fig. 1.... In general, dopants present in the polysilicon layer may diffuse within that layer or into adjacent insulating layers, or into the silicon substrate.

Dopant diffusion in the polysilicon film is poorly characterized at present, although it is known to be much more rapid than in single crystal

silicon. This presumably reflects enhanced diffusion along grain boundaries. Because of a lack of data, specific models, and because it is a reasonable approximation in most cases of interest, SUPREM III will initially model diffusion within the polysilicon film as occurring instantaneously (i.e. $D \simeq \infty$). Thus dopants are assumed to be uniformly distributed in the polysilicon films. More acurate models will be incorporated in updated versions as data and physical models are developed in future work. Our recent results in this area are described in Section 15 of this report.

Segregation of dopants at polysilicon/insulator (SiO_2) interfaces will is handled in SUPREM III exactly as in the case of single crystal/insulator (SiO_2) interfaces. This is also an approximation but is the best that can be done at present because of the lack of specific experimental results and physical models.

This leaves the problem of the polysilicon/single crystal Si interface. We have been engaged in a study of this interface for some time both because of its practical importance (for example in buried contacts in MOS processes) and because of its interest from a materials and modeling point of view. Dopant segregation across this interface, and effects such as OED when the upper polysilicon surface is subjected to thermal oxidation are of interest, and models suitable for incorporation in SUPREM III are the objective.

It is to be expected that any OED effects observed in the polysilicon/single crystal structure will be different from that of single crystal silicon. It has been reported that in such multilayer structures, the tail region of phosphorus diffusion profiles is less pronounced. This seems to indicate that the polysilicon acts as a sink for any vacancies present. In fact, this property of polysilicon has been used to getter defects in the Si substrate. It is believed that Si interstitials generated at the oxidizing polysilicon interface can interact and be absorbed by the grain boundaries when they

diffuse through the polysilicon layer. However, the extent of such inter-action has not been quantified. In this work, the OED effect from the poly-silicon/Si structure and the extent of interaction between the grain boundar-ies and interstitials were determined experimentally and a simple model has been proposed for inclusion in SUPREM III.

The structure used in the experiments is that shown in Figure 1.31. A thin oxide layer, either 150 Å or 300 Å thick, was first grown on a <100> oriented, 5-10 Ω-cm, P type substrate. This thin oxide layer was patterned into 1 mm wide stripes. A layer of polysilicon either 3100 Å or 4700 Å thick was deposited at 800°C in an atmospheric pressure epitaxial reactor. The polysilicon was ion implanted with 10^{16} atoms/cm^2 of phosphorus at 60 keV. An oxide/nitride/oxide sandwich was then deposited and patterned, yeild-ing the structure shown in the figure.

These samples were then subjected to heat treatments at 1200°C, 1000°C or 900°C in a wet oxygen ambient. There are four regions in Figure 1.31. In regions 1 and 2, the thin oxide presents a barrier to the movement of dopants from the polysilicon to the substrate whereas in regions 3 and 4, the polysil-icon is depostied directly on the substrate. In regions 1 and 4, the polysil-icon is exposed to the oxidizing ambient. The nitride layer prevents oxida-tion in regions 2 and 3.

The junction depths in the samples were measured using a staining tech-nique. Concentration profiles were measured using spreading resistance. The proximity of the four regions to one another, made it convenient to compare the effects of oxidation and of the thin oxide barrier, on the diffusion of dopants from the polysilicon to the substrate.

Figure 1.32 shows sample angles lapped on a 34 minute block and stained with a CuSO$_4$, HF solution to reveal the n-type regions. These samples of

3100 A polysilicon on 150 A oxide stripes had been in a wet oxygen ambient at 1200°C for 24 minutes and 36 minutes respectively. During the oxidation three-fourths of the polysilicon exposed to the ambient was consumed for the sample oxidized for 24 minutes. For the sample oxidized for 36 minutes, all the poly-silicon was just consumed. In figure 1.32a, it is seen that where there is the 150 A oxide layer betweeen the polysilcon and the substrate, there is no stain-ing, indicating that the oxide has been effective as a barrier to the diffusion of phosphorus to the substrate. In figure 1.32b, phosphorus is detected in the substrate even in the region with the oxide layer but in which the polysilicon has been oxidized. This is because the pile-up of the phosphorus in front of the oxidizing interface causes a rise in the concentration of phosphorus in the polysilicon, the phosphorus concentration finally reaching such a high value that the thin oxide is converted to phosphorus glass and the phosphorus diffuses into the substrate by "melt through".

Also in Figures 1.32a and 1.32b, it is seen that where the polysilicon is deposited directly on the substrate, the junction depth is greater in the region which is oxidized than in the region protected by nitride. There are two factors which account for this greater junction depth. First, as the poly-silicon is oxidized, the segregation of phosphorus from the oxide to the poly-silicon causes a rise in the concentration of the phosphorus in the polysili-con, which results in a deeper junction in the substrate. The second and more important factor is oxidation enhanced diffusion.

For OED in the oxidation of single crystal silicon, the diffusion coeffi-cient may be written as described earlier [1.11, 1.81]

$$D = D_{N_2} + D_{OED}$$ (1.134)

94

where

$$D_{OED} = D_{OED_{Single\ crystal}}\left(e^{-Zpoly/L_I}\right) \qquad (1.135)$$

As in the case of single crystal, $D_{OED_{Single\ Crystal}}$ contains a $(dx/dt)^{0.5}$ dependence, and in fact is given by the same formulation used in the single crystal case (i.e. eqn. 1.85). In the case of a polysilicon doping source which is being oxidized, the interstitials generated by the oxidation get absorbed as they pass through the polysilicon. The e^{-Zpoly/L_I} term reflects this, where Z_{poly} is the thickness of the polysilicon and L_I is a characteristic adsorption length for the interstitials in the polysilicon (typically $\simeq$ 1000 A).

Figure 1.33 shows an example of the compariscn of experimental and theoretical results using eqn. (1.135). In this case the diffusion temperature was 1000°C, and the polysilicon film thickness was initially 3100 A. In the unoxidized region the junction depth increases as the square root of time as expected. In the oxidized region, however, the junction depth curve is concave upwards. If OED were not a strong function of time, the junction depth in this region, too, would be linearly dependent on the square root of time. The reasonable agreement between experiment and theory has also been obtained at other temperatures and for other polysilicon film thicknesses, indicating that the single crystal models for OED described earlier can be extended to the present situation, provided that the additional term in eqn. (1.135) is included, to account for interstitial absorption in the polysilicon layer.

1.6. EXTRACTION OF ELECTRICAL PARAMETERS BY SUPREM

A process simulation program such as SUPREM-III is used to relate a set of processing steps to the resulting distribution of impurities within a semi-

conductor device structure. This analysis capability can be used for several purposes, including the evaluation of process models, the determination of physical coefficients, and the calculation of electrical parameters. The last of these is probably the most important, allowing a process or device designer to relate a process sequence to the electrical performance of devices fabricated using the process. The capability can provide the basis for optimizing the performance of individual devices and, ultimately the circuits incorporating these devices.

There are two distinct levels of complexity at which the output of a process simulator can be used to evaluate electrical device characteristics. At the highest level, the most complete set of information for guiding device design can be obtained by performing complete one- or two-dimensional device analysis, using the results of process simulation as a basis for specifying the device structure. Device characteristics such as complete current-voltage curves, terminal capacitances, and breakdown voltages can be determined in this manner. Device simulation programs such as SEDAN [1.114] and GEMINI [1.115] have been designed to perform these types of calculations and are best suited for this purpose. The tight coupling of process and device simulation can aid the design of both large- and small-geometry devices. This mode of design is particularly valuable in evaluating the sensitivity of electrical device characteristics to process variations.

Although a complete characterization of electrical device performance may be necessary to evaluate a nearly complete process design, there are many situations where such detailed analysis is unnecessary and ineffective. Device analysis with a low level of complexity is useful in such a case and can provide sufficient direction in the design of a process. Device characteristics such as sheet resistance, conductivity, and capacitance can be

96

calculated using one-dimensional process simulation output as a basis for simple one-dimensional device analysis. This type of analysis is consistent with the fact that some of these simple parameters can be evaluated during the fabrication process without requiring the fabrication of complicated test devices.

The SUPREM-II and SUPREM-III computer programs both provide for the calculation of electrical parameters at the lower level of complexity discussed above. During the evolution from SUPREM-II to SUPREM-III, several issues have prompted an evaluation of the appropriate form for device analysis which should be provided as part of a process simulation program. The accuracy of the electrical parameters calculated by SUPREM-II has been questioned for some process sequences. For example, the calculation of the threshold voltage for depletion-mode MOSFETs is often substantially inaccurate. During the development of the SUPREM programs, there has occurred an evolution of practical device structures which now requires the simulation of very complicated process sequences. While the SUPREM-II program was capable of simulating a two-layer structure typically consisting of silicon covered by silicon dioxide, the SUPREM-III program has been designed to simulate more sophisticated multi-layer structures. Consequently, the virtually automatic extraction of electrical parameters provided by SUPREM-II may no longer be possible or desirable in SUPREM-III.

1.6.A. Electrical Calculations in SUPREM-II

The SUPREM-II computer program calculates several simple electrical parameters associated with the simulated fabrication process. The sheet resistance is determined for each electrical layer bounded by material interfaces or metallurgical junctions. The threshold voltage for MOSFET devices is calculated by assuming that the simulated impurity profile represents the

97

impurity profile in the channel region. Each of these calculations contains inaccuracies, with the degree of error depending on the particular process.

SUPREM-II calculates the sheet resistance ρ_j of layer j using

$$\rho_j = q \left[\int_{y_j}^{y_{j+1}} \mu(C)N(y)dy \right]^{-1} \tag{1.136}$$

where q is the electronic charge, y_j and y_{j+1} are the vertical boundaries of the layer, μ is the carrier mobility, C is the atomic impurity concentration, and N is the electrically active impurity concentration. Eq. (1.136) is incorrect because $N(y)$ is used in place of the mobile carrier concentration, effectively making the assumption that the layer is charge neutral. Consequently, the sheet resistance is usually underestimated because its calculation ignores depletion regions where the carrier concentration is reduced below $N(y)$. The use of Eq. (1.136) is implemented by SUPREM-II because of its simplicity and the desire to avoid a numerical solution of Poisson's equation.

SUPREM-II calculates the threshold voltage V_T for MOSFET devices using the full depletion approximation. All mobile carriers are assumed to be absent from the region $0 < y < y_d$ between the oxide interface and the depletion edge at $y = y_d$. The threshold voltage is defined by

$$V_T = \phi_{MS} - \frac{qN_f}{C_{ox}} + \psi_s + V_{ox} \tag{1.137}$$

where ϕ_{MS} is the gate electrode metal-semiconductor work function, N_f is the oxide interface fixed charge concentration, and C_{ox} is the oxide capacitance.

The term ψ_s in Eq. (1.137) is the surface potential under strong inversion conditions defined by

$$\psi_s = \mp \frac{kT}{q} \ln \left[\frac{N_{inv}N(y_d)}{n_i^2} \right] \tag{1.138}$$

98

where k is Boltzmann's constant, T is the absolute temperature, n_i is the intrinsic carrier concentration, $N(y_d)$ is the impurity concentration at the depletion edge, and N_{inv} is the mobile carrier concentration used to define the onset of strong inversion. In Eq. (1.138), the plus sign applies for p-type substrates and the minus sign applies for n-type substrates.

The term V_{ox} in Eq. (1.137) is the potential drop across the oxide caused by the electric field due to ionized impurity charge within the depletion region and is defined by

$$V_{ox} = - \frac{q}{C_{ox}} \int_0^{y_d} N(y) dy. \qquad (1.139)$$

The depletion edge location y_d in Eqs. (1,138,1.139) is unknown and can be determined by using the full depletion approximation to solve Poisson's equation

$$\frac{d^2\psi}{dy^2} = - \frac{qN(y)}{\varepsilon_s}. \qquad (1.140)$$

Using the boundary conditions

$$\psi(y)\big|_{y=0} = \psi_s \qquad (1.141a)$$

$$\psi(y)\big|_{y=y_d} = 0 \qquad (1.141b)$$

$$\frac{d\psi(y)}{dy}\bigg|_{y=y_d} = 0 \qquad (1.141c)$$

for the semiconductor region and integrating Eq. (1.140) twice yields

$$\psi_s = -\frac{q}{\varepsilon_s} \left\{ y_d \int_0^{y_d} N(y) dy - \int_0^{y_d} \left[\int_0^y N(x) dx \right] dy \right\} \qquad (1.142)$$

SUPREM-II calculates V_T by solving Eqs. (1.137-1.139) and (1.142) simultaneously. There are several problems associated with this technique for determining threshold voltage. The definition of threshold is applicable mainly to enhancement-mode MOSFETs. There is no provision in SUPREM-II for

determining the threshold characteristics of devices such as depletion-mode and buried channel MOSFETs or JFETs. The calculation of the threshold voltage depends on the depletion approximation, which can introduce substantial errors when the substrate impurity concentration is nonuniform. The depletion approximation is also the cause of errors due to the finite Debye length because ψ_S is normally not substantially larger than the thermal voltage kT/q. The last, and possibly most important, difficulty is the assumption that all mobile carriers are absent from the depletion region. This is only valid for a limited set of device structures. For example, depletion-mode and buried channel MOSFETs typically operate with majority carriers present in an implanted channel region within the substrate. However, SUPREM-II determines the location of the depletion edge below the channel region in the substrate and ignores mobile carriers present in the channel.

1.6.B. Electrical Calculations in SUPREM-III

The SUPREM-III computer program calculates the net charge, conductivity, sheet resistance, and capacitance associated with electrical layers bounded by material interfaces and metallurgical junctions. The calculations can be performed for a zero bias condition, which is the default, or for a bias condition established by specifying the bias independently for each layer. A series of successive bias conditions can be simulated by specifying bias steps applied to one or more layers. The result is an extension of the calculations performed by SUPREM-II, allowing the extraction of electrical parameters as a function of applied bias.

SUPREM-III allows the specification of much more general device structures than are possible with SUPREM-II. Some of the assumptions and simplifications which can be made for SUPREM-II are no longer appropriate. The treatment of complex structures and the desire to avoid the problems

100

associated with the electrical parameter calculations performed by SUPREM-II requires that SUPREM-III use solutions of Poisson's equation to calculate electrical device characteristics. In the general case, the one-dimensional Poisson's equation is

$$\frac{d}{dy}\left(\epsilon(y)\frac{d\psi(y)}{dy}\right) = -q[p(y,\psi) - n(y,\psi) + N_D^+(y,\psi) - N_A^-(y,\psi) + N_I(y)] \quad (1.143)$$

where ϵ is the dielectric permittivity, n and p are the mobile electron and hole concentrations, respectively, N_D^+ and N_A^- are the ionized net electrically active donor and acceptor impurity concentrations, respectively, and N_I is the fixed interface charge concentration. The electron and hole concentrations are given by

$$n(y,\psi) = N_c F_{1/2}\left[\frac{q}{kT}\left(\psi(y) - \phi_{Fn}(y)\right)\right] \quad (1.144a)$$

$$p(y,\psi) = N_v F_{1/2}\left[\frac{q}{kT}\left(\phi_{Fp}(y) - \psi(y) - V_G\right)\right] \quad (1.144b)$$

where N_c and N_v are the effective densities of states in the conduction and valence bands, respectively, V_G is the bandgap potential, ϕ_{Fn} and ϕ_{Fp} are the electron and hole quasi-Fermi potentials, respectively, and $F_{1/2}(\eta)$ is the Fermi-Dirac integral of one-half order.

The solution of Eq. (1.143) is performed subject to boundary conditions imposed on ψ or $d\psi/dy$ at the device surface and in the charge neutral portion of the substrate beyond the active device region. The possibility of multiple dielectric interfaces and semiconductor regions only partially depleted of mobile carriers necessitates the use of a numerical technique to solve Eq. (1.143). This approach offers the advantages of eliminating the depletion approximation and correctly treating the depletion regions in the semiconductor. Thus, the inaccuracies encountered in the electrical calculations performed by SUPREM-II are avoided in SUPREM-III.

The solution of Poisson's equation requires that the electron and hole quasi-Fermi potentials be specified throughout the structure. These potentials are taken to be constant within each electrical layer in the structure. The bias of each layer is established by the quasi-Fermi potential for the majority carrier in that layer. The quasi-Fermi potentials for the minority carriers in each layer can either be explicitly set or can be left to be appropriately chosen based on the values for the majority carriers.

In some cases, an electrode in a structure may be electrically isolated from external bias connections. One example of this situation is the floating polysilicon electrode in an EPROM structure. SUPREM-III can approximately treat this case by allowing an electrical layer to be treated as a neutral dielectric, even though it may consist of semiconductor material.

The most basic electrical parameters calculated by SUPREM-III are the total electron and hole concentrations in each electrical layer and in the entire structure. The electron and hole concentrations for layer j are defined as

$$n_j = \int_{y_j}^{y_{j+1}} n(y)\,dy \qquad (1.145a)$$

$$p_j = \int_{y_j}^{y_{j+1}} p(y)\,dy \qquad (1.145b)$$

where n and p are determined by the one-dimensional solution of Poisson's equation. The total electron and hole concentrations for the structure are obtained by summing n_j and p_j over all layers

$$n_T = \sum_{j=1}^{J} n_j \qquad (1.146a)$$

$$p_T = \sum_{j=1}^{J} p_j \qquad (1.146b)$$

102

where J is the number of layers in the structure. The main application for these parameters is in the determination of the charge storage capacity of one or more layers in a device.

SUPREM-III calculates the conductivities associated with both electrons and holes for each layer and for the entire structure. The conductivities for layer j are defined as

$$\sigma_{nj} = q \int_{y_j}^{y_{j+1}} \mu_n(C)n(y)dy \qquad (1.147a)$$

$$\sigma_{pj} = q \int_{y_j}^{y_{j+1}} \mu_p(C)p(y)dy \qquad (1.147b)$$

where μ_n and μ_p are the electron and hole carrier mobilities, respectively, and C is the atomic impurity concentration. The total conductivities for the structure are obtained by summing σ_{nj} and σ_{pj} over all layers

$$\sigma_{nT} = \sum_{j=1}^{J} \sigma_{nj} \qquad (1.148a)$$

$$\sigma_{pT} = \sum_{j=1}^{J} \sigma_{pj} \qquad (1.148b)$$

These parameters are used primarily to determine linear region current-voltage characteristics and threshold or pinchoff voltages. For example, by calculating σ_n versus gate bias for the channel layer in an n-channel depletion-mode MOSFET it is possible to evaluate the threshold voltage of the device when operating in the linear region. The advantages of these calculations over those provided by SUPREM-II are that conductivity can be calculated as a function of bias and the results include the effect of depletion regions.

103

The sheet resistances associated with electrons and holes are calculated by SUPREM-III for each layer and for the entire structure. These values are determined directly from the conductivities discussed above. For layer j, the sheet resistances are based on Eq. (1.147)

$$p_{nj} = \sigma_{nj}^{-1} \qquad (1.149a)$$

$$p_{pj} = \sigma_{pj}^{-1} \qquad (1.149b)$$

The sheet resistances for the entire structure are based on Eq. (1.148)

$$p_{nT} = \sigma_{nT}^{-1} \qquad (1.150a)$$

$$p_{pT} = \sigma_{pT}^{-1}. \qquad (1.150b)$$

These values of sheet resistance should be much more reliable than the SUPREM-II calculations because the presence of depletion regions is taken into account by SUPREM-III. Thus, the problem of sheet resistance being over-estimated by SUPREM-II should be eliminated.

SUPREM-III provides for the calculation of capacitance, a parameter that is not available from SUPREM-II. Capacitance is defined in general as the change in charge for a given change in applied bias. The generality of the structures which can be treated by SUPREM-III poses problems when an attempt is made to determine a capacitance calculation technique which will be appropriate to all structures. The approach taken is to calculate the capacitance associated with the changes in the total electron and hole concentrations for a change in the bias applied to one layer in the structure. The capacitances for a bias change of ΔV are defined as

$$C_n = \frac{\Delta[q \int_{-\infty}^{\infty} n(y)dy]}{\Delta V} \qquad (1.151a)$$

104

$$C_p = \frac{\Delta[q \int_{-\infty}^{\infty} p(y)dy]}{\Delta V} \qquad (1.151b)$$

These calculations can be applied to the determination of electrode and junction capacitance. A further generalization which has been considered is to allow the specification of the layers contributing to the total electron and hole concentrations which are used to determine the capacitance. For example, in the case of an npn bipolar transistor it may be necessary to evaluate the emitter-base capacitance by calculating the change in the electron concentration within the emitter and base layers only, excluding the effect of changes in this concentration within the collector layer.

1.6.C. Example of SUPREM-III Electrical Calculations

The electrical calculations available from SUPREM-III can be illustrated by means of a simple example. Fig. 1.34 shows the input to SUPREM-III for the simulation of the channel region of an n-channel depletion-mode MOSFET. The first set of input commands accomplishes the implantation of an arsenic channel through a 500A thick gate oxide into a p-type substrate. An n^+ polysilicon layer is deposited to provide for a gate electrode. The substrate impurity distribution resulting from this process sequence is shown in Fig. 1.35.

The second set of input commands in Fig. 1.34 accomplishes the extraction of electrical parameters for the depletion-mode device. The initial bias condition is established with a gate bias of -3 volt and zero bias for the channel and substrate. The gate bias is varied by six equal one volt steps and electrical parameters are calculated for each bias condition. The results of these simulations are shown in Fig. 1.36. The first set of information indicates the locations of the boundaries and the depths for each of the layers in the substrate. The charge, conductivity, and sheet resistance for electrons and holes are shown for each bias condition, for each layer, and for the total structure. The capacitance for electrons and holes associated with a change in gate bias is also shown for each bias condition.

105

To provide a visual representation of the electrical extraction capabilities provided by SUPREM-III, some of the most useful information in Fig. 1.36 has been plotted. Fig. 1.37 shows the electron conductivity of the entire structure as a function of the applied gate bias. This curve exhibits characteristics typical of a depletion-mode device. For comparison, the conductivity calculated by SUPREM-II for the same device is shown by an arrow. This value is relatively close to the conductivity calculated by SUPREM-III for zero gate bias. The algorithms in SUPREM-II are unable to determine the threshold voltage of this device. However, using the curve in Fig. 1.37, a designer could calculate the threshold voltage based on a preferred definition.

Fig. 1.38 shows the electron and hole capacitance associated with a change in gate bias as a function of the applied gate bias. The data is extracted from Fig. 1.36. The electron capacitance represents the gate-to-substrate capacitance, while the hole capacitance represents the gate-to-channel capacitance for the device.

In summary, the electrical parameter calculations available in SUPREM-III represent a significant improvement over the capabilities of SUPREM-II. The inaccuracies present in the SUPREM-II calculations have been eliminated by utilizing a numerical solution of the complete nonlinear Poisson equation. The calculation of capacitance has been added and the calculations of charge, conductivity, and sheet resistance have been generalized to be applicable to a wider range of device structures. Finally, the ability to perform calculations at a sequence of applied bias conditions provides a useful, but not overly complicated, device analysis capability within SUPREM-III.

REFERENCES

[1.1] B. E. Deal and A. S. Grove, "General Relationship for the Thermal Oxidation of Silicon," J. Appl. Phys., 36, 3770 (1965)

[1.2] R. H. Doremus, "Oxidation of Silicon by Water and Oxyton and Diffusion in Fused Silica," J. Phys. Chem., 80, 1773 (1976).

[1.3] P. J. Jorgensen, "Effect of Electric Field on Silicon Oxidation," J. Chem. Phys., 37, 874 (1962).

[1.4] D. O. Rayleigh, "Transport Processes in the Thermal Oxidation of Silicon," J. Electrochem. Soc., 113, 782 (1966).

[1.5] F. C. Collins and T. Nakayama, "Transport Processes in the Thermal Growth of Metal and Semiconductor Oxide Films," J. Electrochem. Soc., 114, 1962 (1967).

[1.6] T. G. Mills and F. A. Kroger, "Electrical Conduction at Elevated Temperatures in Thermally Grown Silicon Dioxide Films," J. Electrochem. Soc., 120, 1582 (1973).

[1.7] C. P. Ho and J. D. Plummer, "Si/SiO$_2$ Interface Oxidation Kinetics: A Physical Model for the Influence of High Substrate Doping Levels. I. Theory," J. Electrochem. Soc., 126, 1516-1522 (Sept 1979).

[1.8] C. P. Ho and J. D. Plummer, "Si/SiO$_2$ Interface Oxidation Kinetics: A Physical Model for the Influence of High Substrate Doping Levels. II. Comparison with Experiment and Discussion," J. Electrochem. Soc., 126, 1523-1530 (Sept 1979).

[1.9] P. S. Dobson, "The Effect of Oxidation on Anomalous Diffusion in Silicon," Philosophical Mag., 24, 567-576 (1971).

[1.10] P. S. Dobson, "The Mechanism of Impurity Diffusion in Silicon," Philosophical Mag., 26, 1301-1306 (1972).

[1.11] S. M. Hu, "Formation of Stacking Faults and Enhanced Diffusion in the Oxidation of Silicon," J. Appl. Phys., 45, 1567 (1974).

[1.12] D. A. Antoniadis, A. G. Gonzalez, and R. W. Dutton, "Boron in Near Intrinsic (100) and (111) Silicon under Inert and Oxidizing Ambients--Diffusion and Segregation," J. Electrochem. Soc., 125, 813 (1978).

[1.13] A. M. Lin, D. A. Antoniadis, R. W. Dutton, and W. A. Tiller, "The Rate Control Model of Oxidation-Stacking-Faults Growth in Silicon," ECS Meeting, Oct 1979, Los Angeles, Abstract No. 539.

[1.14] B. E. Deal, "The Current Understanding of Charges in the Thermally Oxidized Silicon Structure," J. Electrochem. Soc., 121, 1980 (1974).

107

[1.15]   J. D. Meindl et al., "Computer-Aided Engineering of Semiconductor Inte-
         grated Circuits," Stanford Electronics Laboratories, TR DXG501 (Jul 1980).

[1.16]   J. D. Plummer et. al., "Computer-Aided Design of Integrated Circuit
         Fabrication Processes for VLSI Devices," Stanford Electronics Laboratories,
         TR DXG501-81 (Jul 1981).

[1.17]   D. W. Hess and B. E. Deal, "Kinetics of the Thermal Oxidation of Silicon
         in O_2/HCl Mixtures," J. Electrochem. Soc., 124, May 1977, pp. 735-739.

[1.18]   B. E. Deal, "Thermal Oxidation Kinetics of Silicon in Pyrogenic H_2O and
         5% HCl/H_2O Mixtures," J. Electrochem. Soc., 125, Apr 1978, pp. 576-579.

[1.19]   R. R. Razouk, L. N. Lie, and B. E. Deal, "Kinetics of High Pressure
         Oxidation of Silicon in Pyrogenic Steam," J. Electrochem. Soc., accepted
         for publication.

[1.20]   B. E. Deal, L. N. Lie, and R. R. Razouk, "Characterization of Silicon
         Oxidation in High Pressure Dry O_2," Extended Abstracts of Fall Meeting
         of the Electrochemical Society, Denver, Colorado, Oct 11-16, 1981,
         Abs. 372.

[1.21]   B. E. Deal and M. Sklar, "Thermal Oxidation of Heavily Doped Silicon,"
         J. Electrochem. Soc., 112, Apr 1965, pp. 430-435.

[1.22]   C. P. Ho, J. D. Plummer, B. E. Deal, and J. D. Meindl, "Thermal Oxidation
         of Heavily Doped Phosphorus Doped Silicon," J. Electrochem. Soc., 125,
         665-671.

[1.23]   W. Schockley and J. L. Moll, "Solubility of Flaws in Heavily-Doped Semi-
         conductors," Phys. Rev., 119, Sept 1960, pp. 1480-1482.

[1.24]   E. A. Irene and D. W. Doug, "Silicon Oxidation Studies: The Oxidation
         of Heavily B- and P-Doped Single Crystal Silicon," J. Electrochem. Soc.,
         125, July 1978, pp. 1146-1151.

[1.25]   R. J. Kriegler, Y. C. Cheng, and D. R. Colton, J. Electrochem. Soc., 119,
         388 (1972).

[1.26]   P. H. Robinson and F. P. Heiman, J. Electrochem. Soc., 118, 141 (1971).

[1.27]   C. M. Osburn, J. Electrochem. Soc., 121, 809 (1974).

[1.28]   K. Hirabayashi and J. Iwamura, J. Electrochem. Soc., 120, 1595 (1973).

[1.29]   Y. J. van der Meulen, C. M. Osburn, and J. F. Ziegler, J. Electrochem.
         Soc., 122, 284 (1975).

[1.30]   D. W. Hess and B. E. Deal, J. Electrochem. Soc., 124, 735 (1977).

[1.31]   G. J. DeClerck, et al., J. Electrochem. Soc., 124, 436 (1975).

[1.32]   E. J. Janssens and G. J. DeClerck, J. Electrochem. Soc., 125, 1696 (1978).

[1.33]   J. D. Meindl, et al., "Final Report on Computer-Aided Semiconductor Process Modeling," Stanford Electronics Laboratories, TR-4969-73F, Oct. 1976.

[1.34]   J. D. Meindl, et al., "Computer-Aided Engineering of Semiconductor Integrated Circuits," Stanford Electronics Laboratories, TR 4969-3, Feb 1978.

[1.35]   B. E. Deal, et al., J. Electrochem. Soc., 125, 339 (1978).

[1.36]   B. E. Deal, J. Electrochem. Soc., 125, 576 (1978).

[1.37]   P. M. Ritchey, Master's Thesis, The Pennsylvania State University, Aug 1976.

[1.38]   Y. J. van der Meulen, "Kinetics of Thermal Growth of Ultra-Thin Layers of SiO_2 on Silicon. Part 1. Experiment," J. Electrochem Soc. 119, 1972, p. 530.

[1.39]   R. Ghez and Y. J. van der Meulen, "Kinetics of Thermal Growth of Ultra-Thin Layers of SiO_2 on Silicon. Part 2. Theory," J. Electrochem. Soc., 119, 1972, p. 1100.

[1.40]   F. P. Fehlner, "Formation of Ultra-Thin Oxide Films on Silicon," J. Electrochem. Soc., 119, 1972, p. 1723.

[1.41]   J. A. Aboaf, "Formation of 20-25 Å Thermal Oxide Films on Silicon at 950-1140°C," J. Electrochem. Soc., 118, 1971, p. 1370.

[1.42]   E. Taft and L. Cordes, "Optical Evidence for a Silicon-Silicon Oxide Interlayer," J. Electrochem. Soc., 126, 1, 1979, p. 131.

[1.43]   K. Hamano, "Breakdown Characteristics in Thin $SiO2$ Films," Japan J. Appl. Phys., 13, 1974, p. 1085.

[1.44]   E. Harari, "Dielectric Breakdown in Electrically Stressed Thin Films of Thermal SiO_2," J. Appl. Phys., 49, 1978, p. 2478.

[1.45]   J. Blanc, "A Revised Model for the Oxidation of Si by Oxygen," Appl. Phys. Lett., 33, 5, 1978, p. 424.

[1.46]   A. E. Irene, "Silicon Oxidation Studies: Some Aspects of the Initial Oxidation Regime," J. Electrochem. Soc., 125, 1978, p. 1708.

[1.47]   A. S. Grove, Physics and Technology of Semiconductor Devices, John Wiley and Sons, New York, 1967.

[1.48]   P. J. Jorgensen, "Effect of Electric Field on Oxidation," J. Chem. Phys., 37. 1962, p. 874.

[1.49]   J. Lindhard, M. Scharff, and M. Schiott, Mat. Fys. Medd. Dan. Vid. Sclsk., 33, 1 (1963).

[1.50]   J. Gibbons and S. Mylroie, Appl. Phys. Lett., 22, 568, Jun 1973.

[1.51] D. A. Antoniadis, S. E. Hansen, and R. W. Dutton, "SUPREM II - A Program for IC Process Modeling and Simulation," Stanford Electronics Laboratories, Stanford University, TR SEL78-020 (Jun 1978).

[1.52] J. W. Mayer, L. Erikson, and J. A. Davies, Ion Implantation in Semiconductors, New York: Academic Press (1974).

[1.53] W. K. Hofker, D. P. Oosthoek, N. J. Koelman, and H. A. M. De Grefte, Rad. Effects, 24, 223 (1975).

[1.54] W. K. Hofker, H. W. Werner, D. P. Oosthoek, and H. A. M. De Grefte, Appl. Phys., 2, 265, 1973.

[1.55] H. Ryssel, H. Kranz, K. Muller, R. A. Henkelmann, and J. Biersack, Appl. Phys., Lett., 30, 399 (Apr 1977).

[1.56] H. Ryssel, private communication.

[1.57] K. B. Winterbon, Rad. Effects, 13, 215 (1972)

[1.58] D. K. Brice, Rad. Effects, 6, 77 (1970).

[1.59] J. D. Meindl, et al., "Computer-Aided Engineering of Semiconductor Integrated Circuits," Stanford Electronics Laboratories, TR 77-26874-3 (Jul 1979).

[1.60] L. A. Christel, J. F. Gibbons, and S. Mylroie, "Application of the Boltzmann Transport Equation to Ion Range and Damage Distributions in Multilayered Targets," submitted to J. Appl. Phys.

[1.61] J. Lindhard, V. Nielsen, M. Scharff, Mat. Fys. Medd. Dan. Vid. Selsk., 36, No. 14 (1970).

[1.62] K. B. Winterbon, P. Sigmund, and J. B. Sanders, Mat. Fys. Medd. Dan. Vid. Selsk., 37 No. 14 (1970).

[1.63] S. Kalbitzer and H. Oetzmann, Proc. of Intern. Conf. on Ion Beam Modification of Materials (1978).

[1.64] W. D. Wilson, L. G. Haggmark, J. P. Biersack, Phys. Rev., B-15, No. 5, 2458 (1977).

[1.65] T. Hirao, K. Inoue, S. Takayanagi, Y. Yaegashi, J. Appl. Phys., 50, 1 (1979).

[1.66] A. Goetzberger, D. J. Bartelink, J. P. McVittie, and J. F. Gibbons, Appl. Phys. Lett., 29, 259, 8 (1976).

[1.67] R. B. Fair, in Proc. 3rd Intl. Symp. on Silicon Materials Science and Technology (77-2), J. Electrochem. Soc., May 1977, p. 968.

[1.68] R. B. Fair, "Boron Diffusion in Silicon-Concentration and Orientation Dependence, Background Effects, and Profile Estimation," J. Electrochem. Soc., 122, June 1975, p. 800.

[1.69] R. B. Fair, "Impurity Doping Processes in Silicon," edited by F. F. Y. Wang (North Holland Publishing, Amsterdam), in press.

[1.70] R. B. Fair and G. R. Weber, "Effect of Complex Formation in Diffusion of Arsenic in Silicon," J. Appl. Phys., 44, Jan 1973, p. 273

[1.71] J. Murota, et al., "Relationship Between Total As and Electrically Active As in Silicon Produced by the Diffusion Process," J. Appl. Phys., 50, Feb 1979, p. 804.

[1.72] R. O. Schwenker, E. S. Pan, and R. F. Lever, "Arsenic Clustering in Silicon," J. Appl. Phys., 42, July 1971, p. 3195.

[1.73] R. B. Fair and J. C. C. Tsai, "A Quantitative Model for the Diffusion of Phosphorus in Silicon and the Emitter Dip Effect," J. Electrochem. Soc., 124, July 1977, p. 1107.

[1.74] R. B. Fair, "The Effect of Strain-Induced Band-gap Narrowing on High Concentration Phosphorus Diffusion in Silicon," J. Appl. Phys., 50, Feb 1979, p. 860.

[1.75] R. B. Fair, "Oxidation, Impurity Diffusion, and Defect Growth in Silicon - An Overview," J. Electrochem. Soc., 128, June 1981, p. 1360.

[1.76] R. B. Fair, private communication.

[1.77] J. A. van Vechten and C. D. Thurmond, "Entropy of Ionization and Temperature Variation of Ionization Levels of Defects in Semiconductors," Phys. Rev. B, 14, Oct 1976, p. 3539.

[1.78] S. Matsumoto, et al., "The Distribution of the Excess Vacancies in the Bulk at the Diffusion of Phosphorus into Silicon," Jap. J. Appl. Phys., 16, July 1977, p. 1177.

[1.79] M. Yoshida, et. al., "Excess Vacancy Generation Mechanism at Phosphorus Diffusion into Silicon," J. Appl. Phys., 45, Apr 1974, p. 1498.

[1.80] D. A. Antoniadis, A. M. Lin, and R. W. Dutton, "Oxidation-Enhanced Diffusion of Arsenic and Phosphorus in Near-Intrinsic (100) Silicon," Appl. Phys. Lett., 33, Dec 1978, p. 1030.

[1.81] K. Taniguchi, K. Kurosawa, and M. Kashiwagi, "Oxidation Enhanced Diffusion of Boron and Phosphorus in (100) Silicon," J. Electrochem. Soc., 127, Oct 1980, p. 2243.

[1.82] S. Mizuo and H. Higuchi, "Retardation of Sb Diffusion in Si During Thermal Oxidation," Jap. J. Appl. Phys., 20, Apr 1981, p. 739.

[1.83] I. R. Sanders and P. S. Dobson, "Oxidation, Defects and Vacancy Diffusion in Silicon," Philosphical Mag., 20, 1969, p. 881.

[1.84] S. M. Hu, "Formation of Stacking Faults and Enhanced Diffusion in the Oxidation of Silicon," J. Appl. Phys., 45, Apr 1974, p. 1567.

[1.85] S. M. Hu, "Atomic Diffusion in Semiconductors," edited by D. Shaw (Plenum, New York, 1973), Chap. 5.

[1.86] S. P. Murarka, "Role of Point Defects in the Growth of the Oxidation-Induced Stacking Faults in Silicon, II. Retrogrowth, Effect of HCl Oxidation and Orientation," Phys. Rev. B, 21, Jan 1980, p. 692.

[1.87] A. M. Lin, D. A. Antoniadis, and R. W. Dutton, "The Oxidation Rate Dependence of Oxidation-Enhanced Diffusion of Boron and Phosphorus in Silicon," J. Electrochem. Soc., 128, May 1981, p. 1131.

[1.88] A. M. Lin, et al., "The Growth of Oxidation Stacking Faults and the Point Defect Generation at Si-SiO$_2$ Interface during Thermal Oxidation of Silicon," J. Electrochem. Soc., 128, May 1981, p. 1121.

[1.89] S. M. Hu, "Oxygen, Oxidation Stacking Faults, and Related Phenomena in Silicon," Presented at Materials Research Soc. Meeting, Boston, Nov 17-20, 1980.

[1.90] Y. Ishikawa, et. al., "The Enhanced Diffusion of Arsenic and Phosphorus in Silicon by Thermal Oxidation," to be published in J. Electrochem. Soc.

[1.91] D. A. Antoniadis, A. G. Gonzalez, and R. W. Dutton, "Boron in Near-Intrinsic (100) and (111) Silicon under Inert and Oxidizing Ambients - Diffusion and Segregation," J. Electrochem. Soc., 125, May 1978, p. 813.

[1.92] Y. Nabeta, et al., "Restrained Diffusion of Boron and Phosphorus in Silicon under HCl-Added Oxygen Atmosphere," J. Electrochem. Soc., 123, Sept 1976, p. 1416.

[1.93] R. Reif and R. W. Dutton, "Computer Simulation in Silicon Epitaxy," J. Electrochem. Soc., Vol. 128, No. 4, pp. 909-918, April 1981.

[1.94] R. Reif, T. I. Kamins, and K. C. Saraswat, J. Electrochem. Soc., Vol. 126, p. 653 (1979).

[1.95] R. Reif, T. I. Kamins and K. C. Saraswat, J. Electrochem. Soc., Vol. 125, p. 1860 (1978).

[1.96] R. Reif, T. I. Kamins and K. C. Saraswat, J. Electrochem. Soc., Vol 126, p. 644 (1979).

[1.97] D. C. D'Avanzo, Stanford Electronics Labs, Stanford University, Tech. Rep. SEL 78-011, Feb 1978.

[1.98] W. C. Johnson and P. T. Panausis, IEEE Trans. Elec. Dev., ED-18, p. 965 (1971).

[1.99] K. T. Aust and J. W. Rutter, "Recovery and Recrystallization of Metals", Edited by L. Himmel, Interscience Publishers, p. 131 (1962).

[1.100] D. Gupth, D. R. Campbell and P. S. Ho, "Thin Films--Interdiffusion and Reaction", Wiley-Interscience Publication, p. 161 (1980).

[1.101] C. P. Ho and J. D. Plummer, J. Electrochem. Soc., 126, p. 1516 (1979).

[1.102] M. M. Mandurah, K. C. Saraswat, C. R. Helms and T. I. Kamins, Dopant Segregation in Polycrystalline Silicon," J. Appl. Phys., 51, 5755, (1980).

[1.103] J. W. Y. Seto, "The Electrical Properties of Polycrystalline Silicon," J. Appl. Phys., 467, 5247, (1975).

[1.104] M. M. Mandurah, K. C. Saraswat and T. I. Kamins, "A Model for Conduction in Polycrystalline Silicon," IEEE Trans. Electron Dev., ED-33, Oct 1981.

[1.105] E. A. Irene, E. Tierney and D. W. Dong, "Silicon Oxidation Studies: Morphological Aspects of the Oxidation of Polycrystalline Silicon," J. Electrochem. Soc., 127, 705 (1980).

[1.106] T. I. Kamins and E. L. MacKenna, Metal Trans., 2, 2292 (1971).

[1.107] H. Sunami, "Thermal Oxidation of Phosphorus-Doped Polycrystalline Silicon in Wet Oxygen," J. Electrochem. Soc., 125, 892 (1978).

[1.108] T. I. Kamins, "Oxidation of Phosphorus-Doped Low Pressure and Atmospheric Pressure CVD Polycrystalline-Silicon Films," J. Electrochem. Soc., 126, 838 (1979).

[1.109] J. J. Barnes, J. M. DeBlasi and B. E. Deal, "Low Temperature Differential Oxidation for Double Poly-Silicon VLSI Device," J. Electrochem. Soc., 126, 1779 (1979).

[1.110] L. Baldi, C. Ferla and M. Toshi, "Oxidation Kinetics of As-Doped Polysilicon in Steam Environment," Abs. No. 168, Extended Abstracts of the 157th Meeting of Electrochem. Soc., Vol. 80-1, St. Louis, May, 1980.

[1.111] G. Baccarini, B. Ricco and G. Spadini, J. Appl. Phys., 49, 1978, p. 5565.

[1.112] N. C. C. Lu, L. Berzberg and J. D. Meindl, IEEE Elec. Dev. Letters., EDL-1, 1980.

[1.113] L. Mei, B. Swaminathan, R. W. Dutton, "Process Modeling of Multilayer Structures Involving Polycrystalline Silicon", 1980 IEEE IEDM Conference, p. 219-222.

[1.114] D. C. D'Avanzo, M. Vanzi, and R. W. Dutton, "One-Dimensional Semiconductor Device Analysis (SEDAN)," Stanford Electronics Laboratory Technical Report No. G-201-5, Oct. 1979.

[1.115] J. A. Greenfield, S. E. Hansen, and R. W. Dutton, "Two-Dimensional Analysis for Device Modeling," Stanford Electronics Laboratory Technical Report No. G-201-7, July 1980.
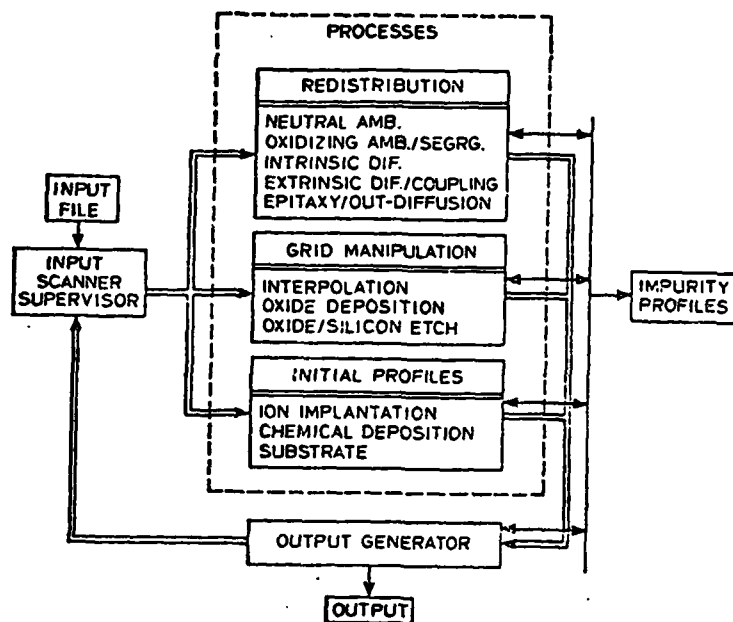
Fig. 1.1: Basic structure of SUPREM Process Modeling Program.
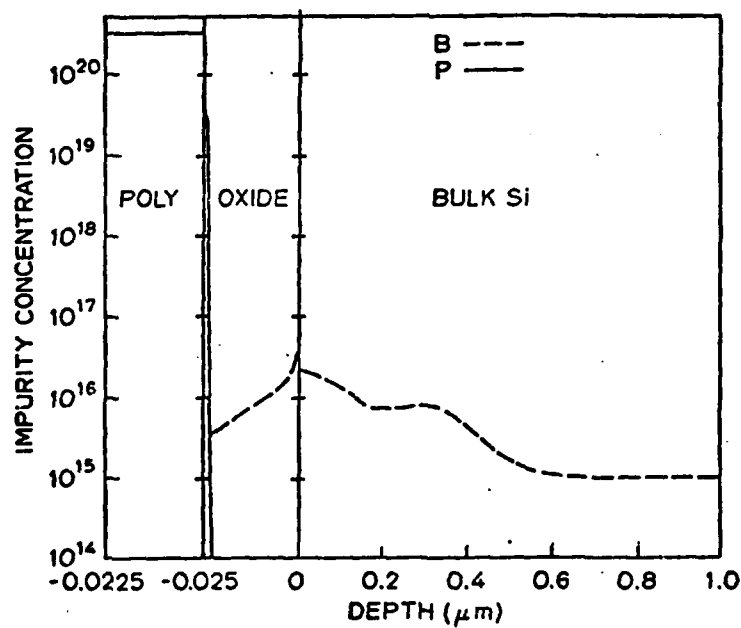
Fig. 1.2: Multi-layer impurity profile simulation for early version of SUPREM III.

114
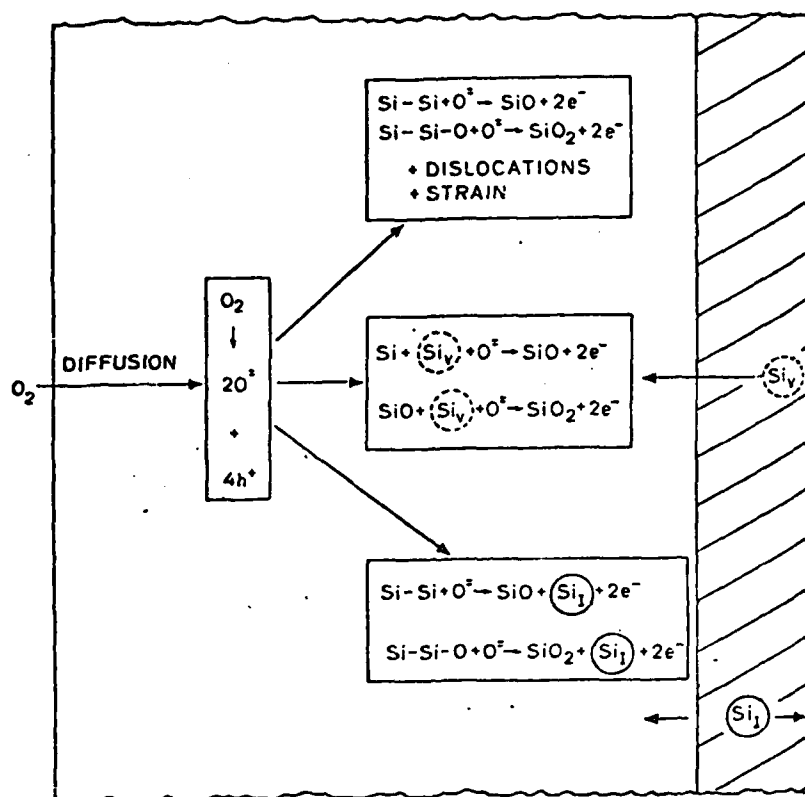
Fig. 1.3: Detailed model of the interface oxidation reaction including possible sources for the required "free volume" for the reaction $Si + O_2 \rightarrow SiO_2$ to proceed
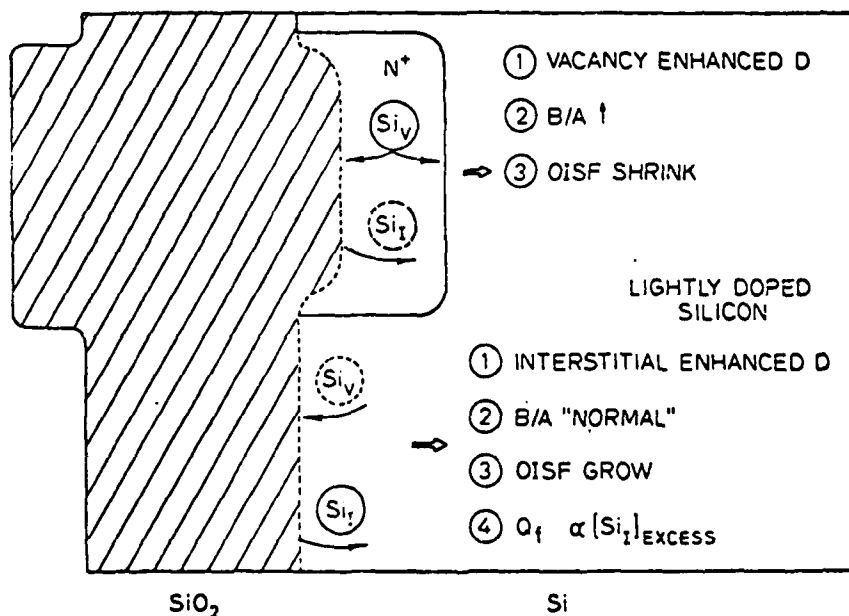
Fig. 1.4: Qualitative effects of interstitial generation and vacancy consumption during thermal oxidation.

115

Fig. 1.6: Normalized parabolic rate constant $\frac{B(P,T)}{B(1,T)}$ vs. O_2 pressure for (100) and (111) silicon oxidized in dry O_2 at 1, 5, 10, 15, and 20 atm. Temperature range was 800°C - 1000°C.
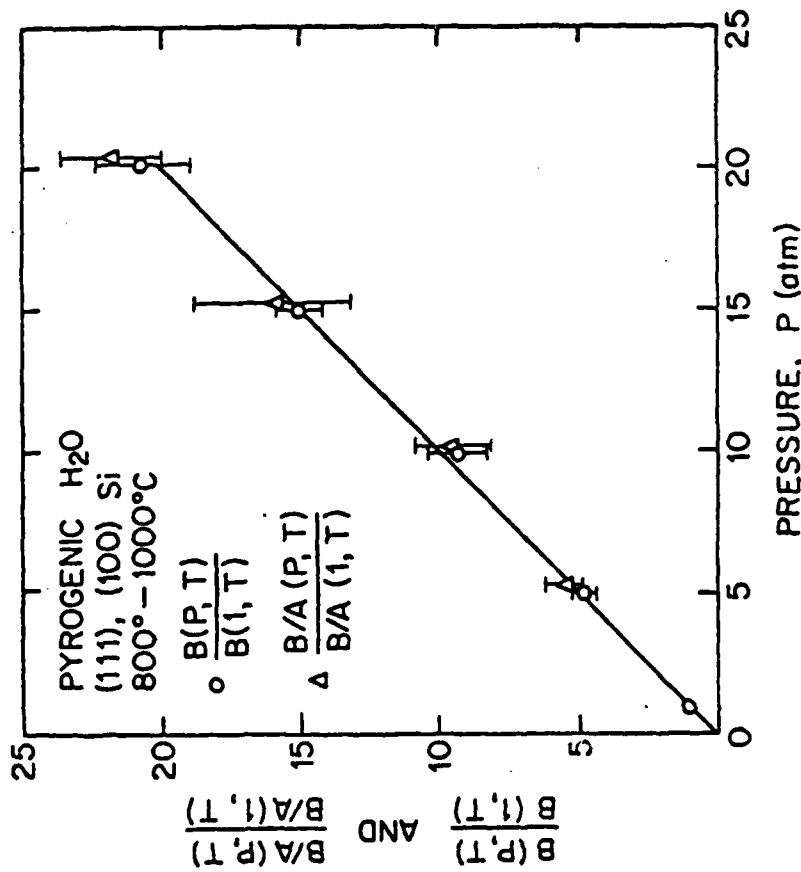
Fig. 1.5: Normalized parabolic and linear rate constant $\frac{B(P,T)}{B(1,T)}$ and $\frac{B/A(P,T)}{B/A(1,T)}$ vs. H_2O pressure for (100) and (111) silicon oxidation in pyrogenic H_2O ambient at 1, 5, 10, 15, and 20 atm for the temperature range 800°C to 1000°C.
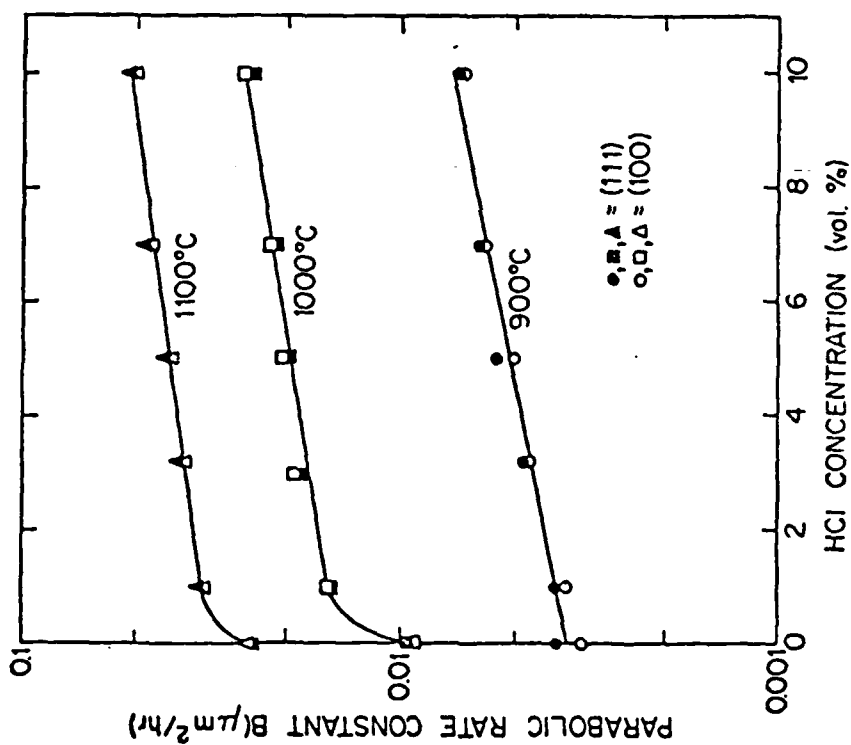
Fig. 1.8: Parabolic Rate Constant vs. %HCl for (111) and (100) oriented N-Type Silicon at 900°C, 1000°C, and 1100°C.
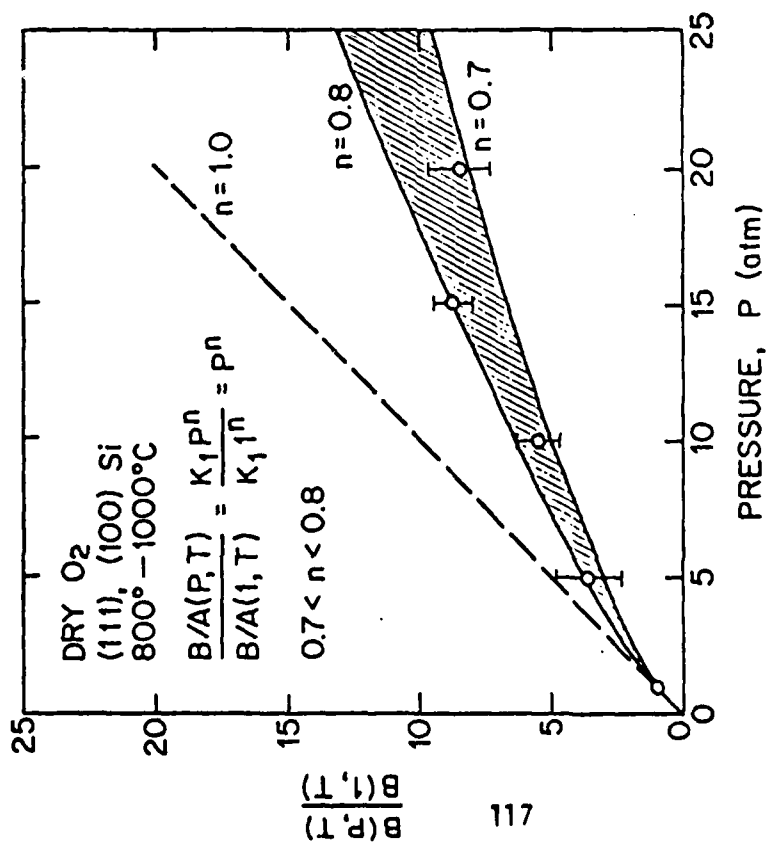
Fig. 1.7: Normalized linear rate constant $\dfrac{B/A(P,T)}{B/A(1,T)}$ vs. O_2 pressure for (100) and (111) silicon oxidized in dry O_2 at 1, 5, 10, 15 and 20 atm. Temperature range was 800°C – 1000°C.

117

Fig.1.10: Oxidation rate dx/dt vs. oxide thickness for (111) and (100) wafers at 800, 850, 900, 950 and 1000°C in dry O2. (P = 1 atm).

Fig. 1.9: Linear Rate Constant vs. %HCl for (111) and (100) Oriented N-Type Silicon at 900°C, 1000°C and 1100°C.

118
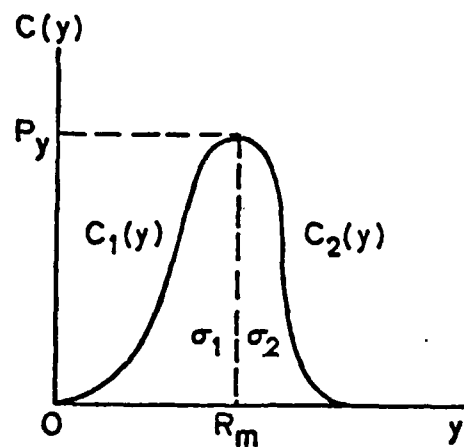
Fig.1.11: The joint half-Gaussian representation of as-implanted impurity profile.
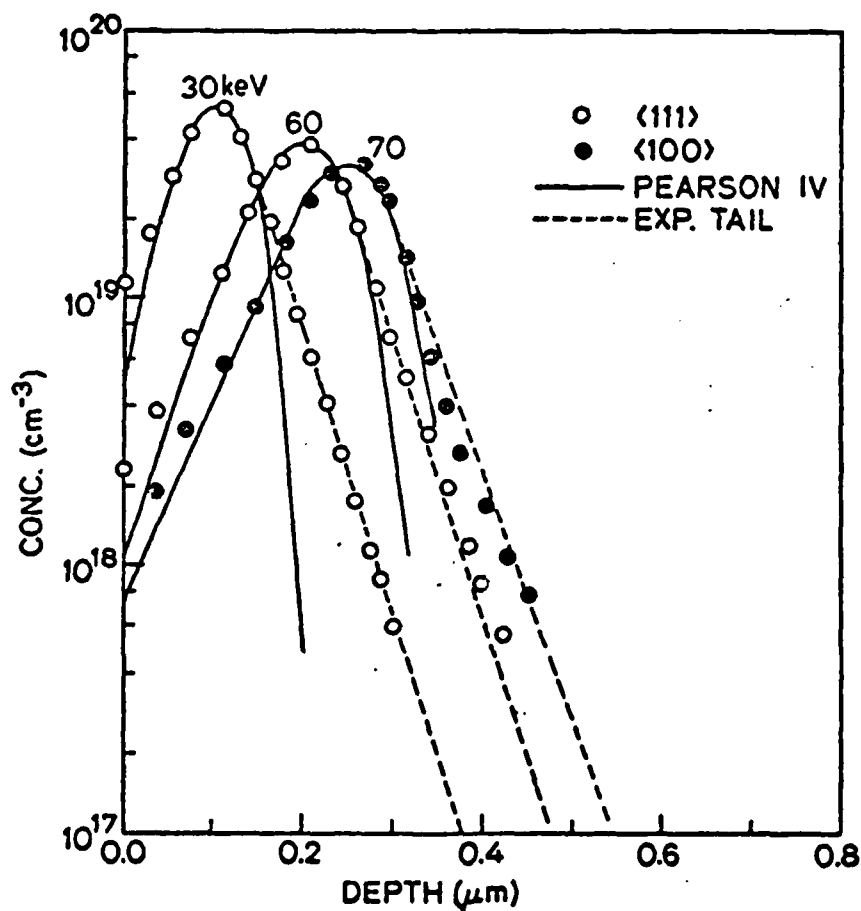
Fig.1.12: Boron as-implanted profiles in (111) and (100) silicon in a random direction. Pearson IV and modified Pearson IV distributions were used. Dose is $10^{13}/cm^2$.

119
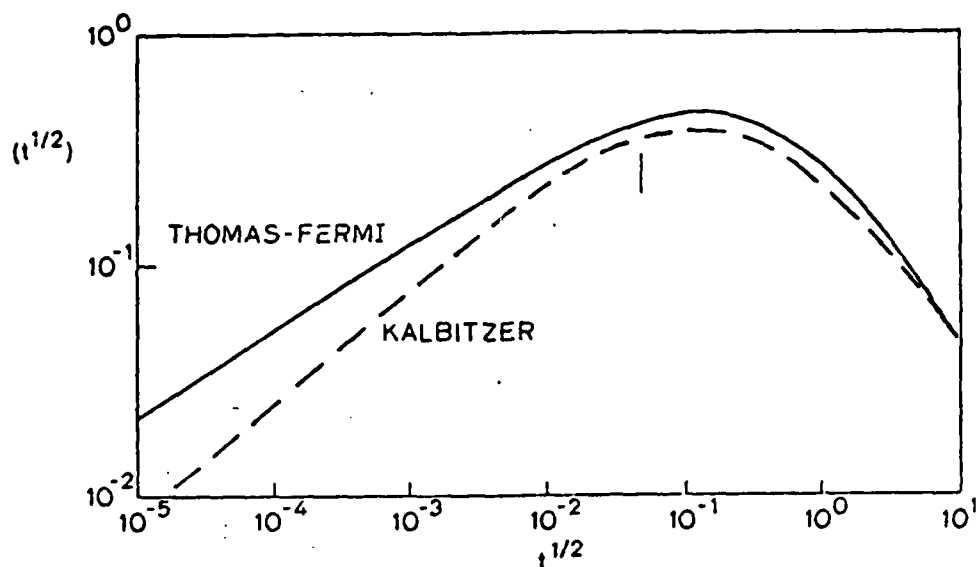
Fig. 1.13: Scattering functions $f(t^{1/2})$ used in the nuclear scattering cross section $d\sigma(t) = 1/2\pi a^2 t^{-3/2} f(t^{1/2})$. The marker indicates the point at which the hybrid cross section switches from the Kalbitzer form on the left to the Thomas Fermi form on the right.
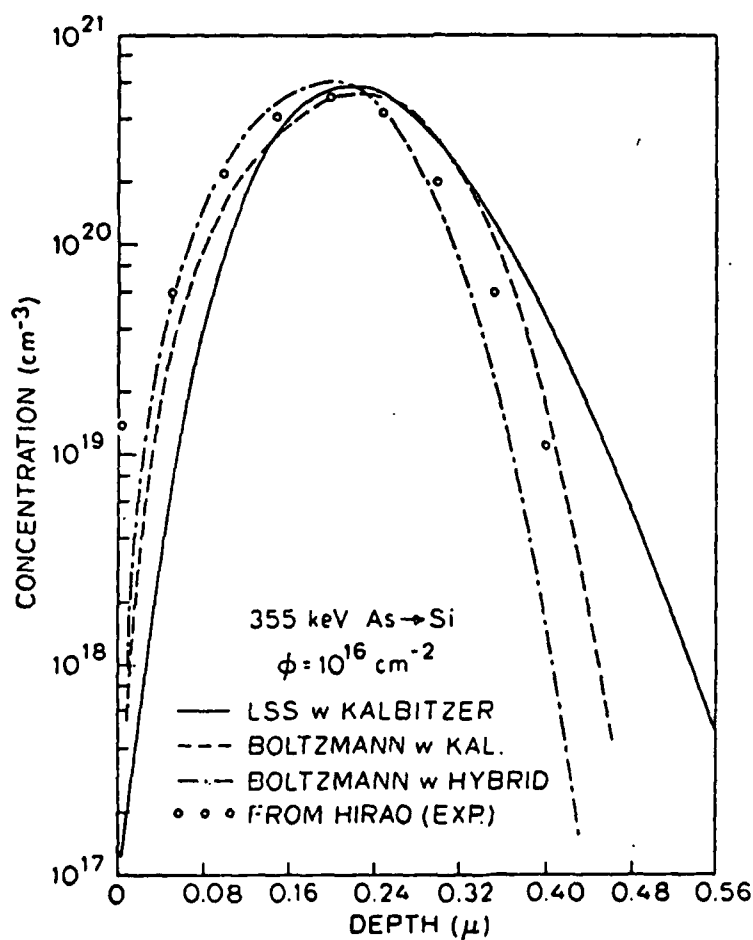
Fig. 1.14: Comparison of LSS and transport equation calculations and experimental results for the range profile of 355 keV arsenic implanted into silicon to a dose of 10^{16} cm^{-2}. The cross section used for each calculation is indicated in the key.

120
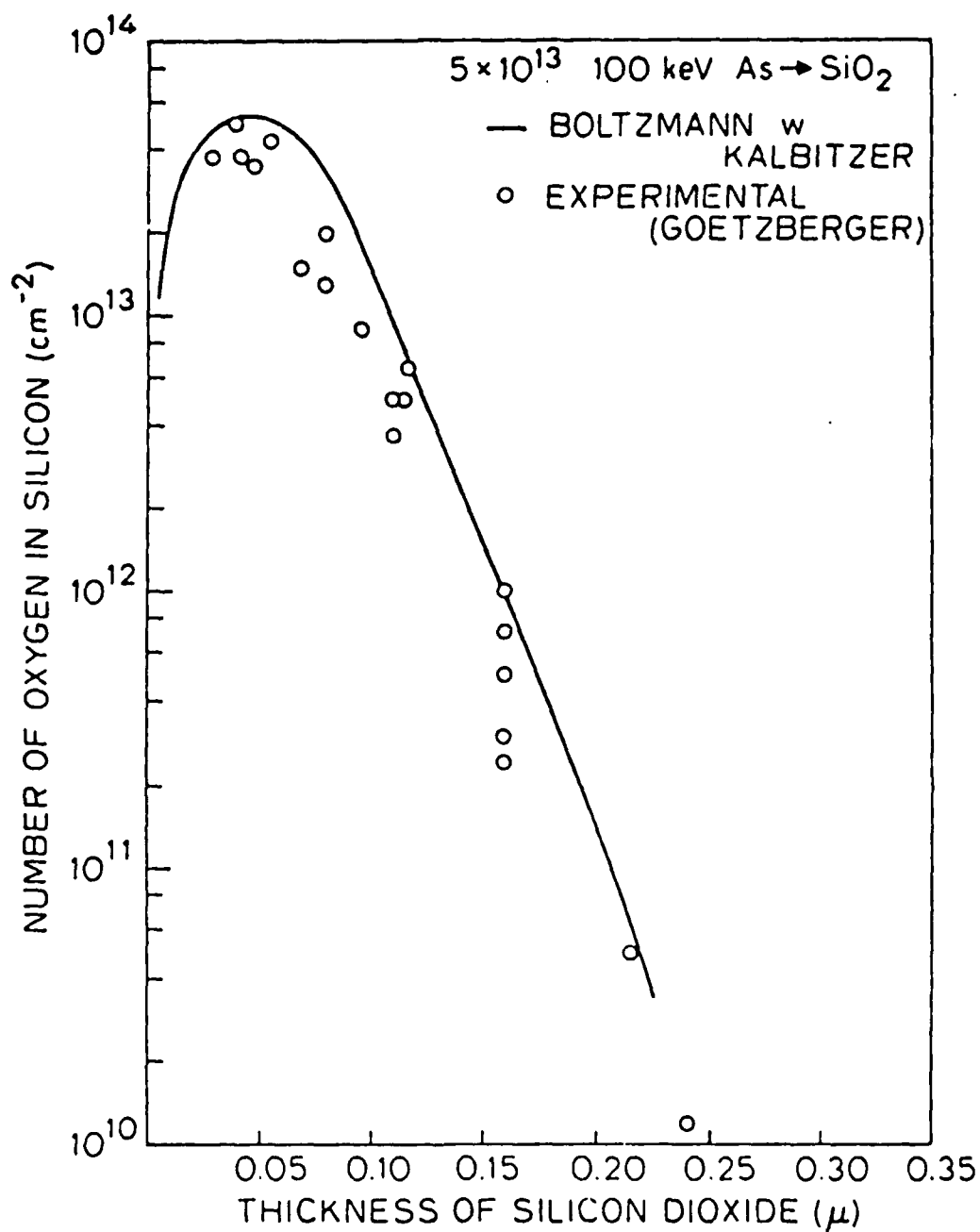
Fig. 1.15: Comparison of a transport equation calculation and experimental results for total oxygen recoil fluxes into silicon resulting from an arsenic implant at 100 keV through various thicknesses of oxide.

```
TITL      SUPREM-2 BIPOLAR EXAMPLE
GRID      YMAX=.7, DPTH=.3, DYSI=.002
SUBS      ELEM=-, CONC=0, ORNT=111

COMM      BASE IMPLANT
STEP      TYPE=IMPL, ELEM=B, DOSE=5.2E13, AKEV=25

COMM      PAD OXIDATION
STEP      TYPE=OXID, TIME=60, TEMP=850, MODL=NITO

COMM      P+ OXIDATION
STEP      TYPE=OXID, TIME=30, TEMP=800, MODL=NITO

COMM      EMITTER IMPLANT
STEP      TYPE=IMPL, ELEM=AS, DOSE=3.0E15, AKEV=40

PLOT      TOTL=Y, AXIS=T, WIND=.7, CMIN=15, NDEC=5.3

COMM      ANNEAL
STEP      TYPE=OXID, TIME=45, TEMP=950, MODL=NITO

SAVE      FILE=ARK04, TYPE=A
END
```

Fig. 1.16a: SUPREM II input code for bipolar transistor example.

```
title       SUPREM-3 Bipolar example

initialize  <111> Si   dX=.007 thickness=.7   phos    conc=3e16

comment     base implant
implant     boron    dose=5.2e13   energy=25

comment     pad oxidation
diffusion   time=60   temp=850

comment     P+ oxidation
diffusion   time=30   temp=800

comment     emitter implant
implant     arsenic  dose=3.0e15   energy=40
print       layer

comment     anneal
diffusion   time=45   temp=950       neutral

plot        total  xmax=.7  cmin=1e15   cmax=1e21  xpwid=6.5

save        name=srk05  structure

stop
```

Fig. 1.16b: SUPREM III input code for bipolar transistor example.
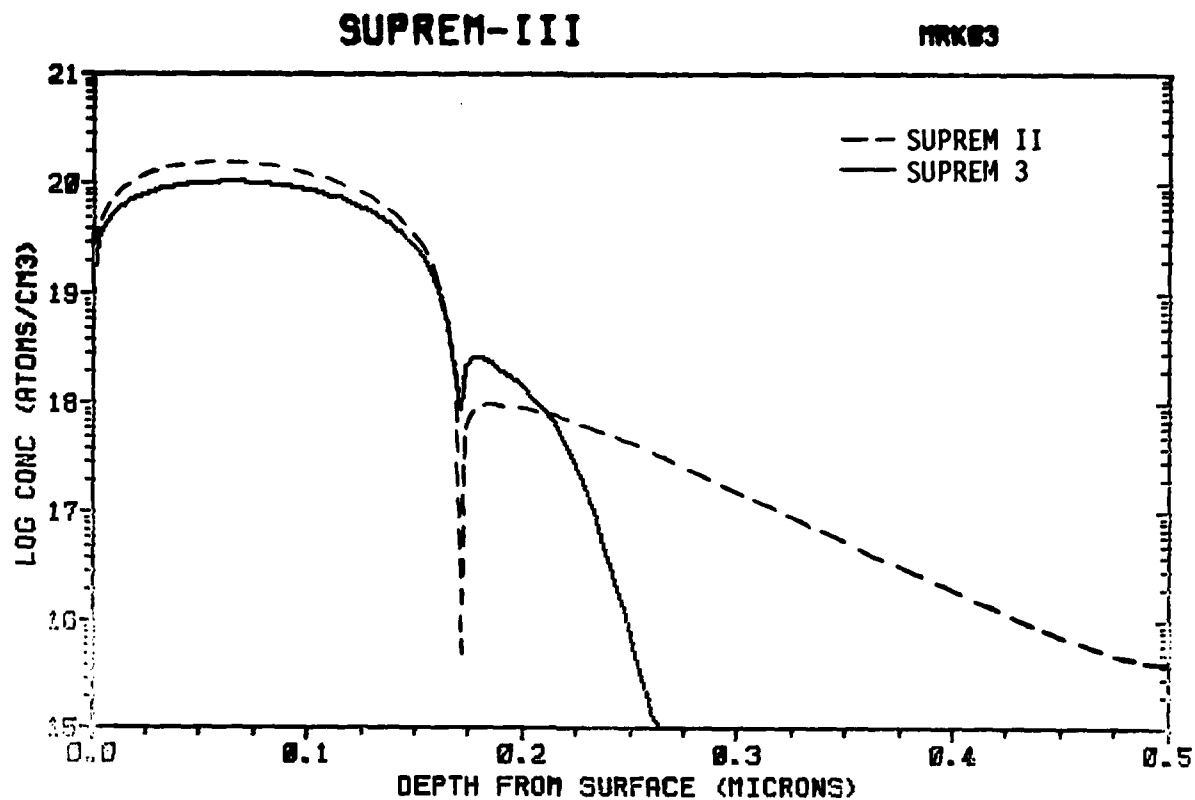
Fig. 1.17: SUPREM II and SUPREM III outputs for the bipolar example of Fig. 1.16.
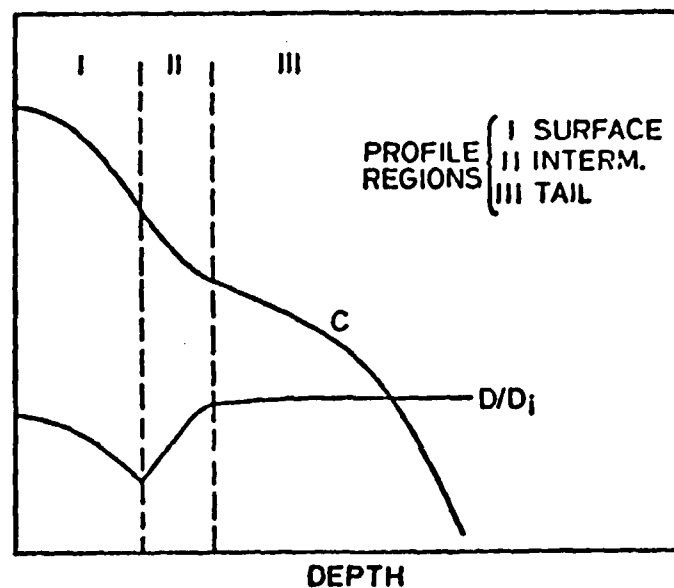
Fig.1.18: A typical phosphorus doping profile
with a demarkation of the three regions
considered by the Fair and Tsai model.
Also shown in the local diffusivity
profile derived by the model.

Fig. 1.20: Phosphorus and arsenic diffusion coefficients in inert and dry O_2 oxidizing ambients as a function of temperature [1.80].

Fig. 1.19: Calculated phosphorus profiles for two different surface concentrations. Slowed diffusion of the higher surface concentration profile is due to the silicon stress model.

GAS   SILICON   ZERO FLUX BOUNDARY

C_g^* C_I

F_s
F_b

MOVING INTERFACE

F_s: EVAPORATION FLUX

F_b: MOVING-BOUNDARY FLUX
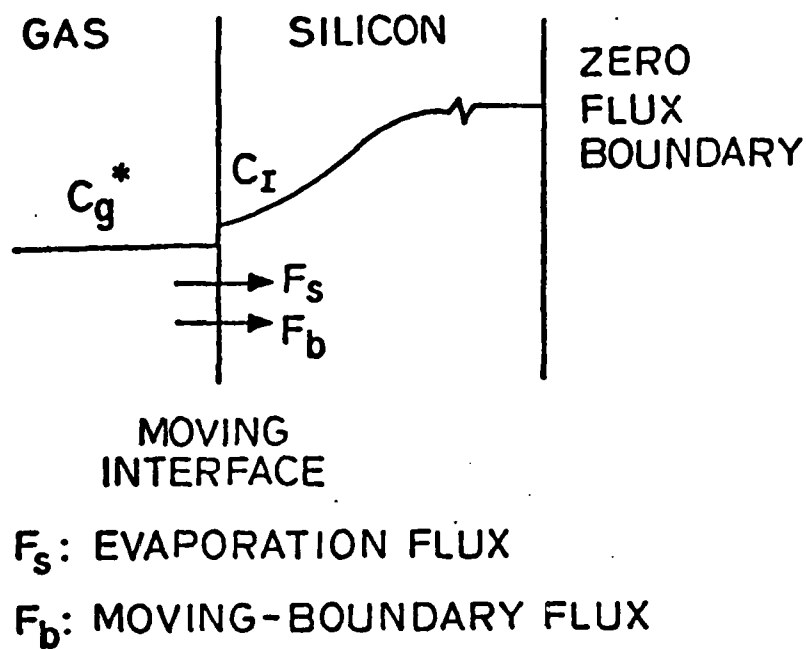
Fig.1.21: Model for the redistribution of impurities during epitaxial silicon growth.

z

GAS PHASE

$z = z_f$ EPI LAYER

$N(z,t)$ SUBSTRATE

$z = \infty$
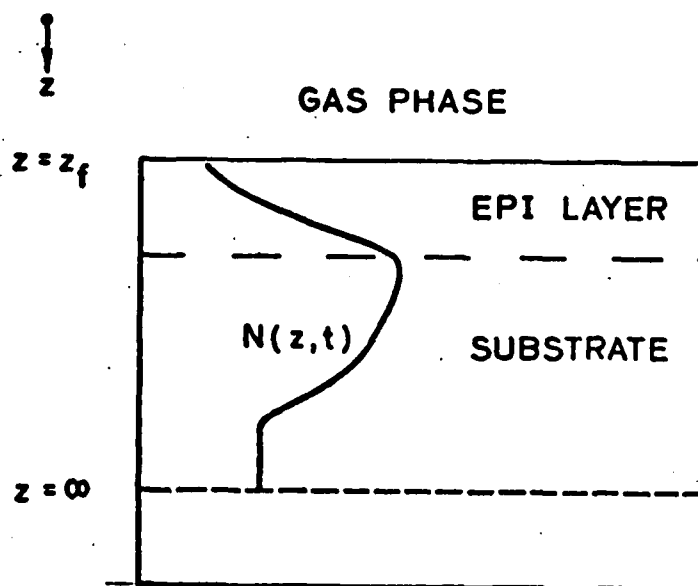
Fig.1.22: Schematic cross-section of a silicon wafer for the purpose of solving Fick's Second Law.

125
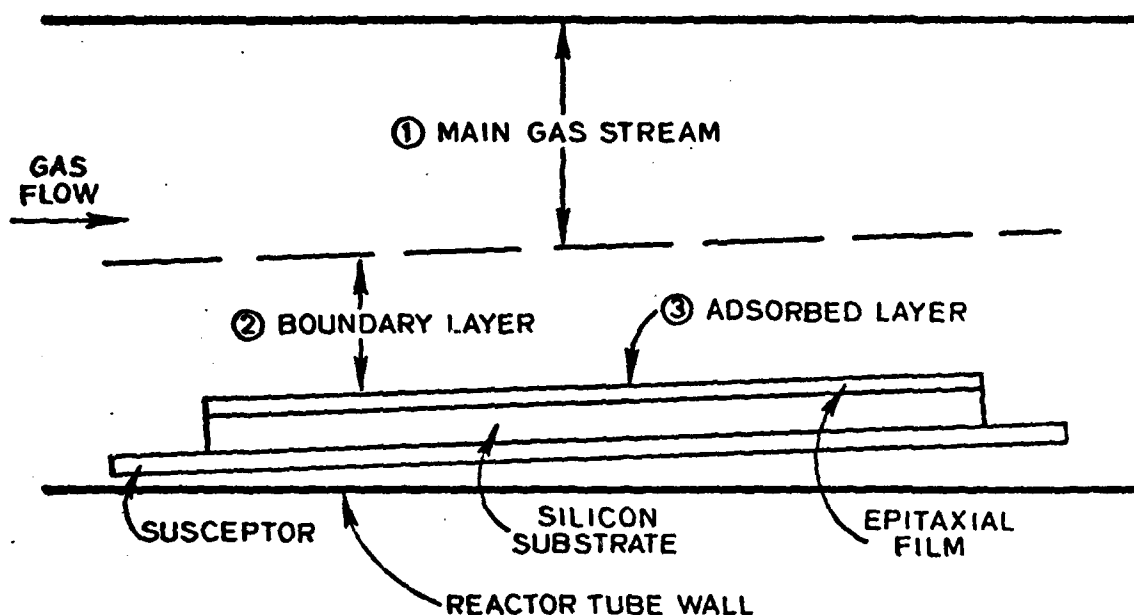
Fig. 1.23: Schematic section of the horizontal reactor tube along the length of the susceptor.
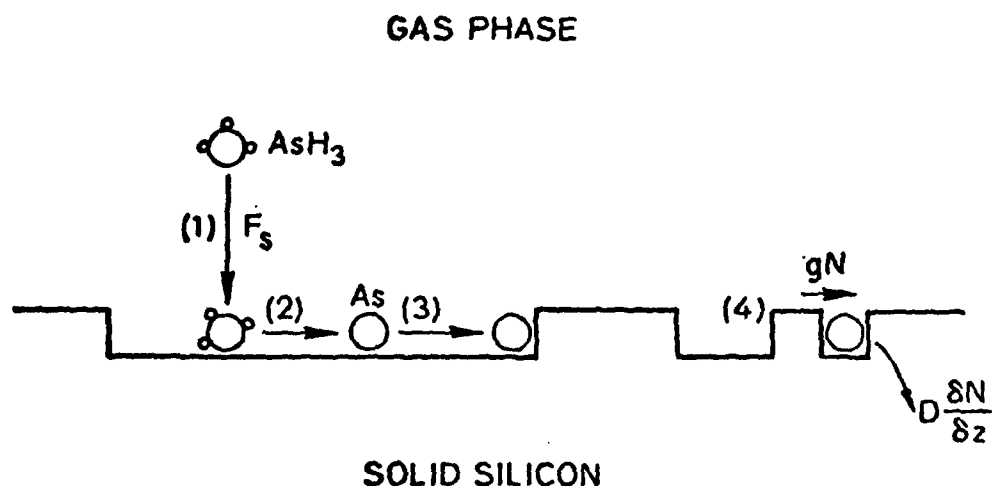
Fig. 1.24: Steps taking part in the doping process and occurring at the silicon surface: (1) arsine adsorption, (2) arsine chemical decomposition, (3) arsine surface diffusion and site incorporation, and (4) covering of incorporated arsenic by subsequently arriving silicon atoms.

126
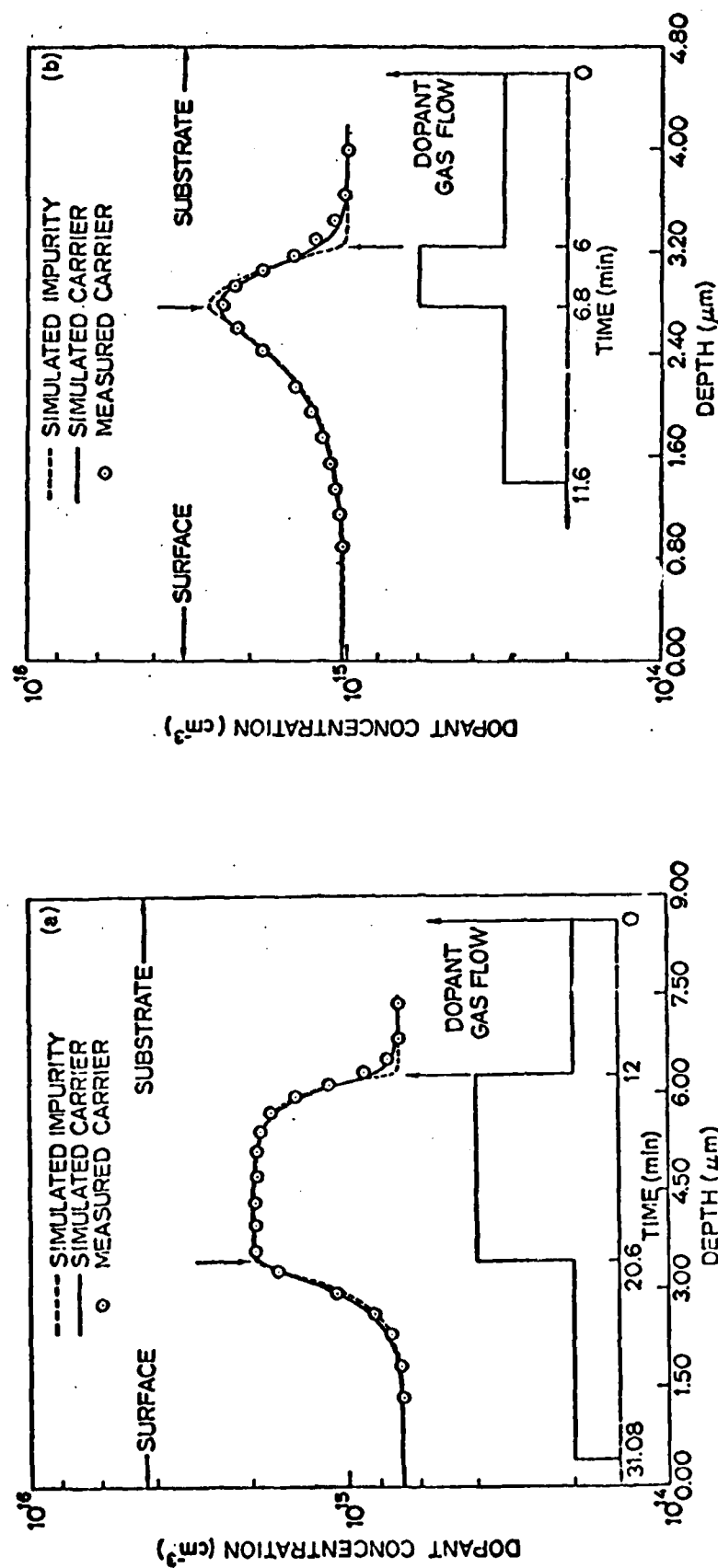
Fig. 1.25: Measured and simulated epitaxial doping profiles corresponding to the time-varying arsine flows indicated in the insets.

127
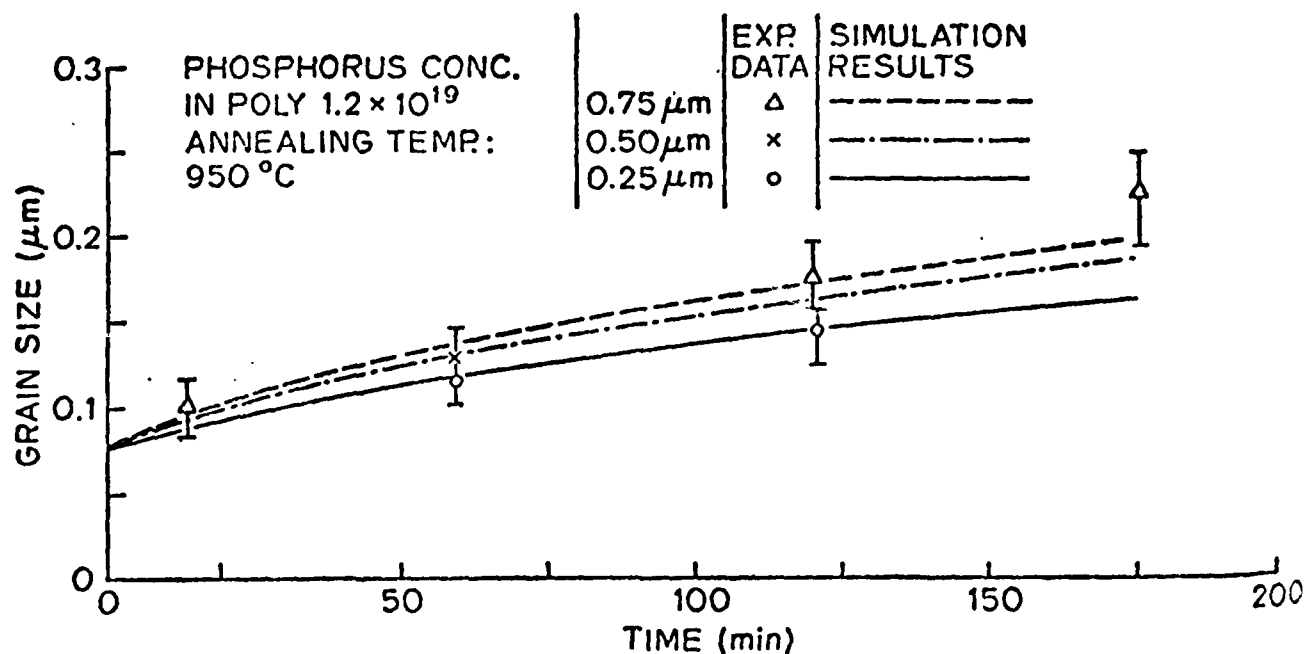
Fig.1.26: Change in average grain size as a function of annealing time at 950°C for three different thicknesses of poly.
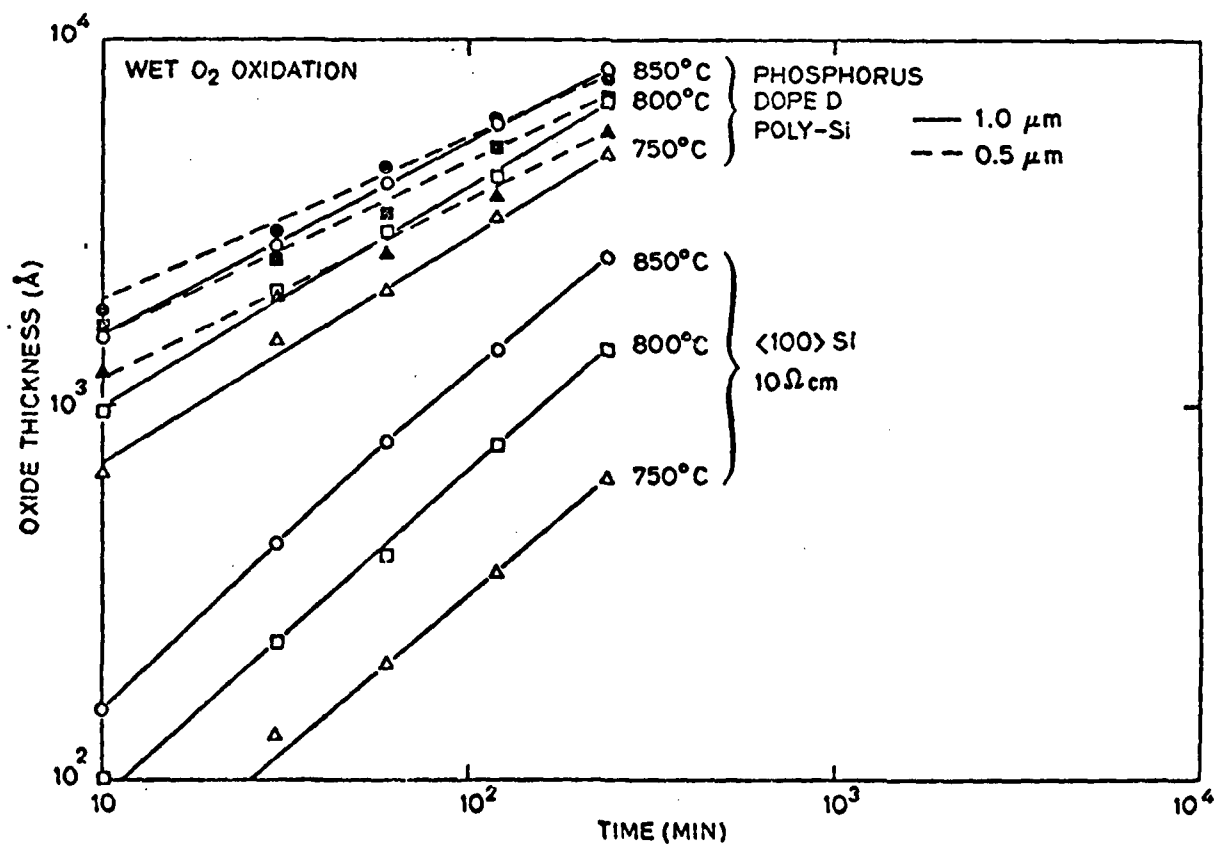
Fig.1.27: Oxide thickness vs. oxidation time for steam (O_2 bubbled through 95°C H_2O) ox dation of phosphorus doped polysilicon films and lightly doped (100) silicon.

Fig.1.29: Model of n-type polycrystalline material. (a) Small crystallites surrounded by grain boundaries containing a large number of traps, (b) resulting space charge, (c) corresponding energy-band diagram with potential barriers surrounding the grain boundaries.

Fig.1.28: Effect of dopant segregation and carrier trapping on the conductivity of polycrystalline-silicon.

129
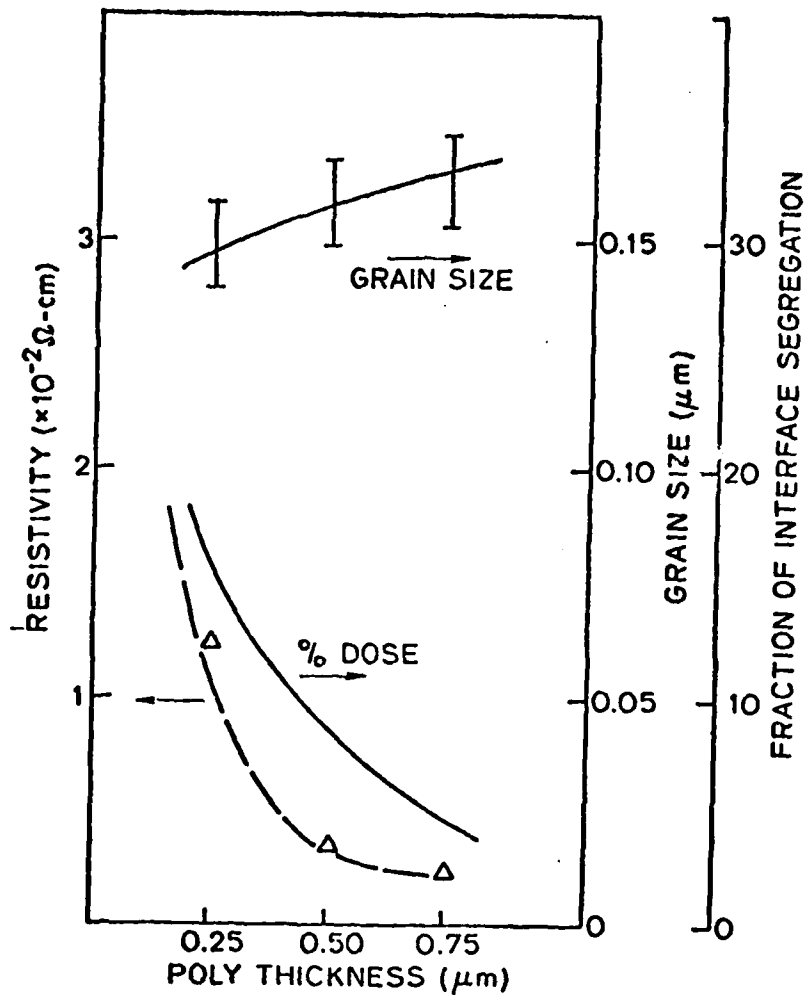
Fig. 1.30: Resistivity, grain size and fraction of dopant
segregated at grain boundaries as functions of
poly thickness for phosphorus doped films annealed
at 950°C for three hours.
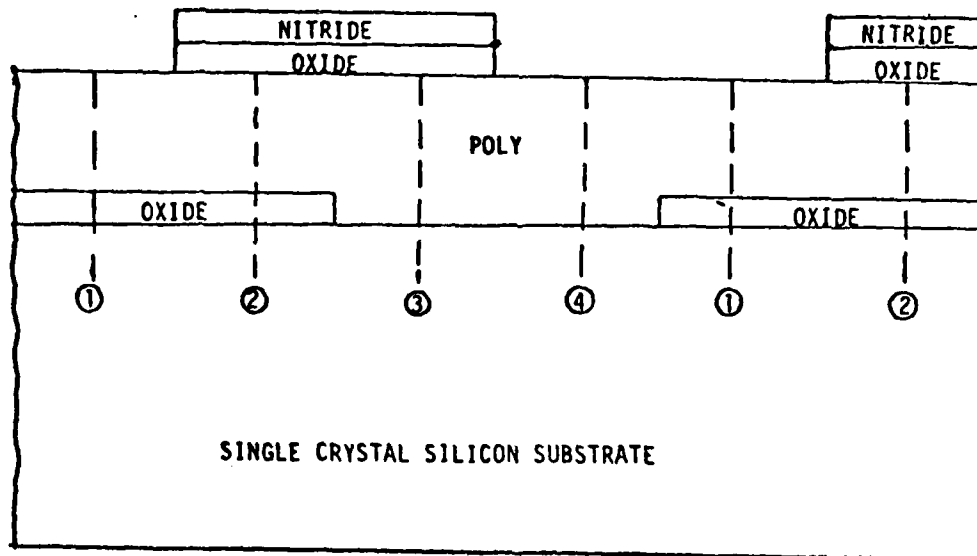
Fig. 1.31:Generalized polysilicon structure along with other
thin film and bulk Si layers of interest for process
modeling.

130
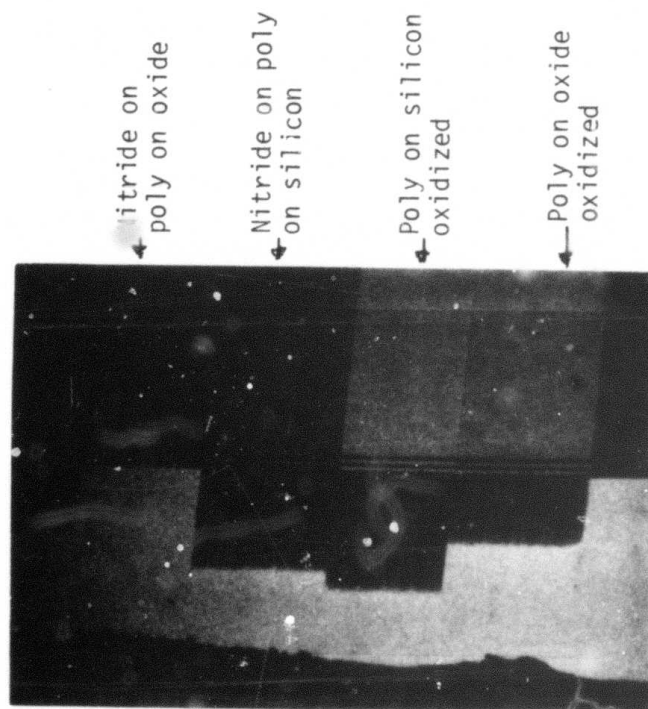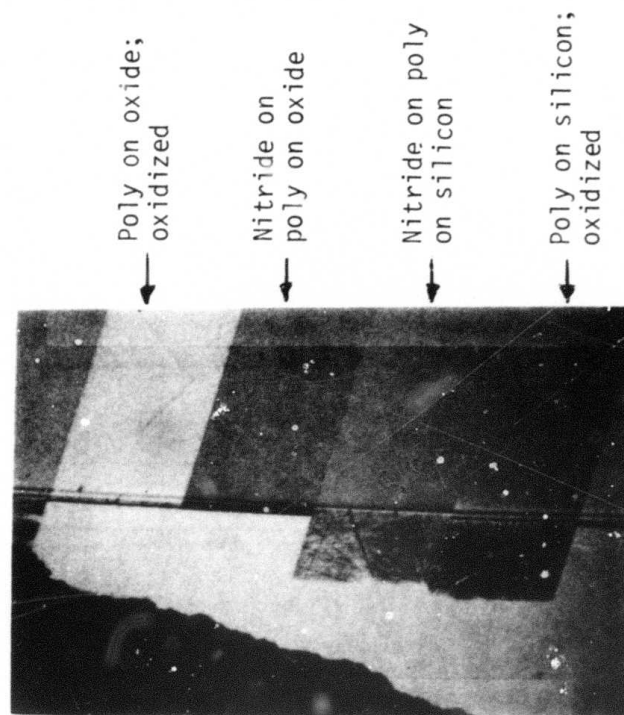
Fig.1.32: OED experiments with polysilicon, with and without 150 Å
SiO$_2$ layer under poly, with and without nitride on poly.
a) Top: 24 min., 1200°C, wet O$_2$. b) Bottom: 36 min.,
1200°C, wet O$_2$.

131

3878A POLY, 1000C
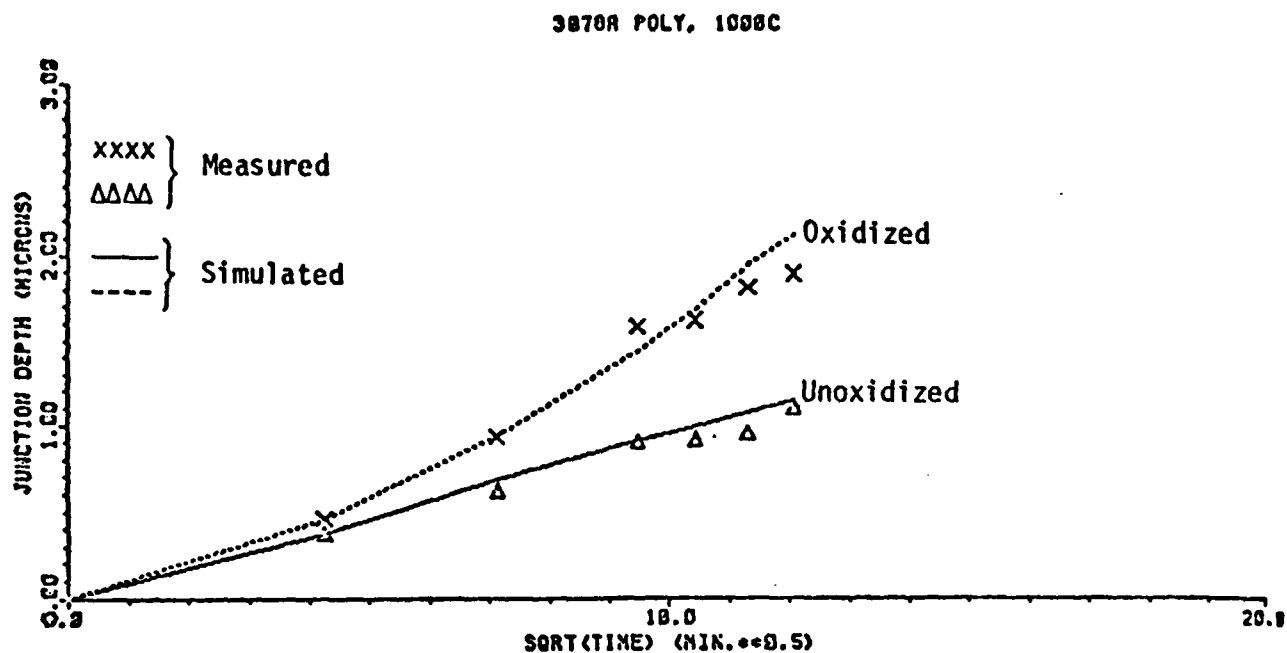
Fig. 1.33: Comparison of Theory (SUPREM) and experiment for junction depth vs. $\sqrt{t}$ for a 3100 Å doped poly layer in oxidized and unoxidized regions (T = 1000°C).

132

```
TITLE   EXAMPLE 2 - DEPLETION-MODE N-CHANNEL MOSFET

INITIALIZE  <100>  BORON  CONCENTRATION=1E15  DEPTH=1  DX=.01
DEPOSIT  OXIDE  THICKNESS=.06
IMPLANT  ARSENIC  ENERGY=100  DOSE=1E12
DEPOSIT  POLY  THICKNESS=.5  PHOSPHORUS  CONCENTRATION=1E18

ELECTRICAL  STEPS=6  DEPTH=2
BIAS          LAYER=3                          V.MAJOR=-3  VDEL.MAJ=1
BIAS          LAYER=1  DIFF.LAY=2  V.MAJOR=0
BIAS          LAYER=1  DIFF.LAY=1  V.MAJOR=0

END
```

Fig.1.34: SUPREM III input simulating the channel region of an n-channel
depletion-mode MOSFET. The input commands are shown for
specification of the process sequence and initiation of the
calculation of electrical parameters.
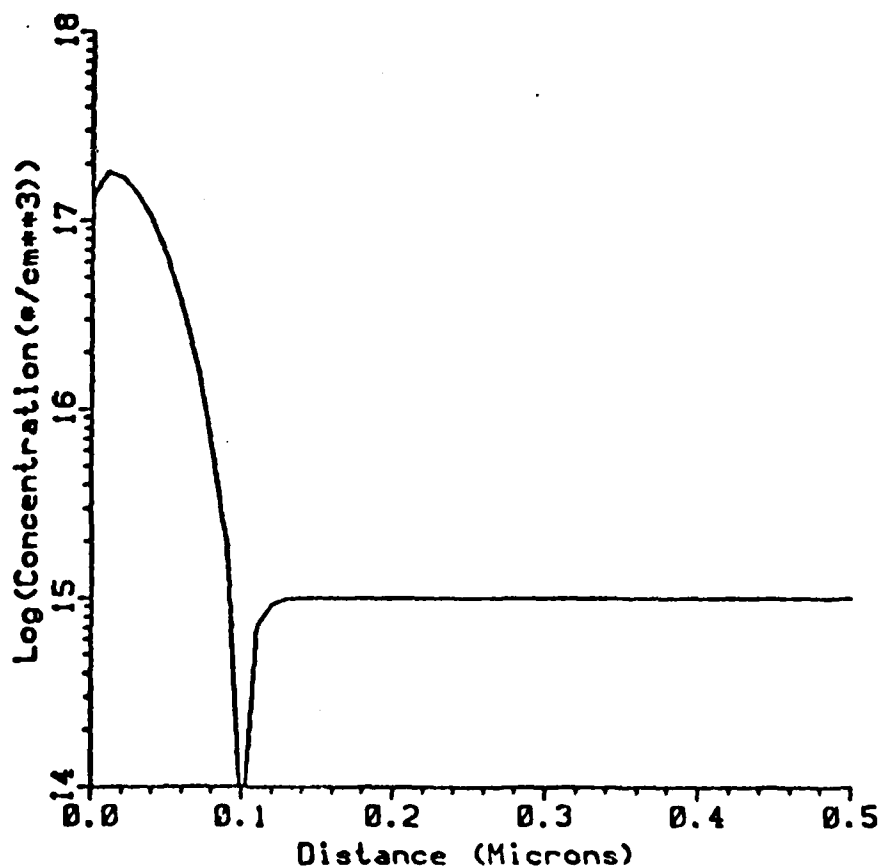
Fig.1.35: The impurity profile in the semiconductor resulting from the
process sequence shown in Fig. 1.34.

133

```
I      Y1          Y2          delta Y
1  0.0000E+00  9.0004E-02  9.0004E-02
2  9.0004E-02  2.0000E+00  1.9100E+00

V(3,1) = -3.0000E+00
I     Total n      Total p      n Cond.      p Cond.      n Rho        p Rho
1  1.7987E+02  1.2999E+11  2.6498E-14  6.4222E-06  3.7738E+13  1.5571E+05
2  1.2741E+02  1.5541E+11  2.6658E-14  1.2227E-05  3.7513E+13  8.1787E+04
T  3.0727E+02  2.8540E+11  5.3156E-14  1.8649E-05  1.8813E+13  5.3622E+04
C  1.6880E-16  6.2820E-08

V(3,1) = -2.0000E+00
I     Total n      Total p      n Cond.      p Cond.      n Rho        p Rho
1  1.8929E+10  1.2747E+02  2.4811E-06  6.2891E-15  4.0305E+05  1.5900E+14
2  1.9737E+09  1.0679E+11  4.1313E-07  8.4019E-06  2.4206E+06  1.1902E+05
T  2.0902E+10  1.0679E+11  2.8942E-06  8.4019E-06  3.4552E+05  1.1902E+05
C  4.0460E-08  1.9360E-09

V(3,1) = -1.0000E+00
I     Total n      Total p      n Cond.      p Cond.      n Rho        p Rho
1  3.4173E+11  0.0000E+00  3.7105E-05  0.0000E+00  2.6951E+04  1.7014E+38
2  9.0021E+09  1.0384E+11  1.8836E-06  8.1697E-06  5.3090E+05  1.2240E+05
T  3.5073E+11  1.0384E+11  3.8988E-05  8.1697E-06  2.5649E+04  1.2240E+05
C  5.9870E-08  4.8000E-11

V(3,1) =  0.0000E+00
I     Total n      Total p      n Cond.      p Cond.      n Rho        p Rho
1  7.3371E+11  0.0000E+00  7.3880E-05  0.0000E+00  1.3536E+04  1.7014E+38
2  9.6641E+09  1.0371E+11  2.0220E-06  8.1593E-06  4.9455E+05  1.2256E+05
T  7.4337E+11  1.0371E+11  7.5902E-05  8.1593E-06  1.3175E+04  1.2256E+05
C  6.5630E-08  1.6000E-11

V(3,1) =  1.0000E+00
I     Total n      Total p      n Cond.      p Cond.      n Rho        p Rho
1  1.1503E+12  0.0000E+00  1.1349E-04  0.0000E+00  8.8112E+03  1.7014E+38
2  9.7919E+09  1.0369E+11  2.0488E-06  8.1574E-06  4.8810E+05  1.2259E+05
T  1.1601E+12  1.0369E+11  1.1554E-04  8.1574E-06  8.6550E+03  1.2259E+05
C  6.7520E-08  1.6000E-11

V(3,1) =  2.0000E+00
I     Total n      Total p      n Cond.      p Cond.      n Rho        p Rho
1  1.6745E+12  0.0000E+00  1.5432E-04  0.0000E+00  6.4799E+03  1.7014E+38
2  9.8448E+09  1.0367E+11  2.0598E-06  8.1565E-06  4.8549E+05  1.2260E+05
T  1.6843E+12  1.0367E+11  1.5638E-04  8.1565E-06  6.3946E+03  1.2260E+05
C  6.7680E-08  1.6000E-11
```

Fig.1.36: The electrical parameters calculated by SUPREM III for the input
sequence shown in Fig.1.34. Charge, conductivity, sheet resistance,
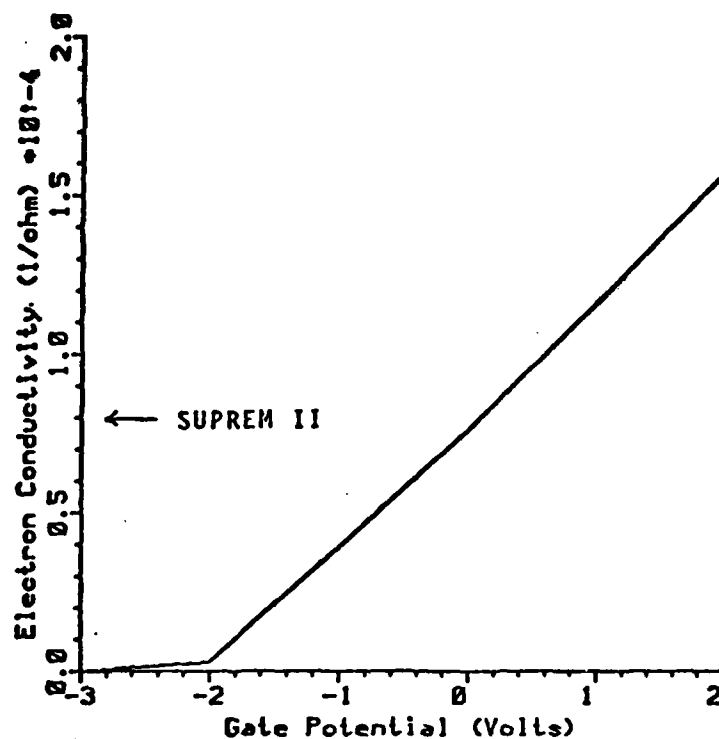and capacitance are shown for each of six bias conditions.

Fig. 1.37: The electron conductivity versus gate bias for the entire substrate of the n-channel depletion-mode device represented by the process sequence shown in Fig. 1.34.
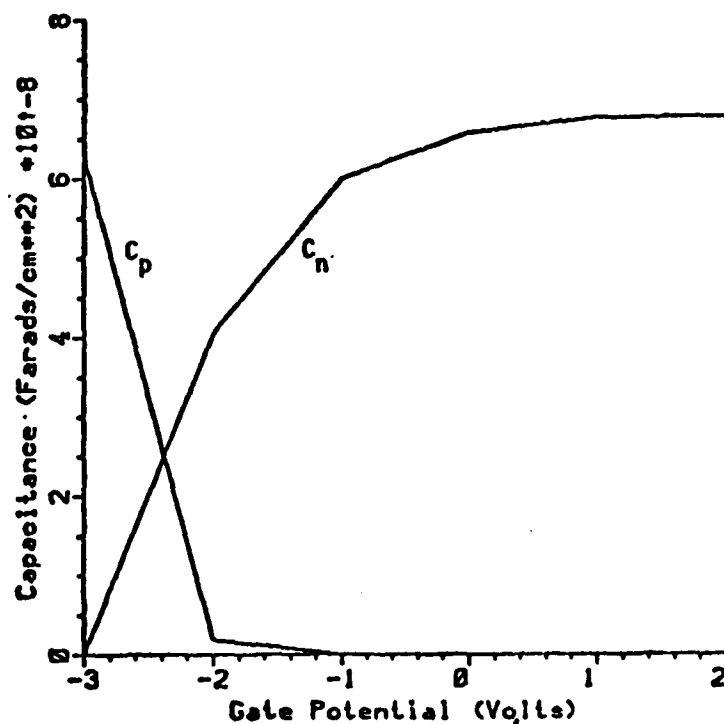

Fig. 1.38: The electron and hole capacitance versus gate bias for the n-channel depletion-mode device represented by the process sequence shown in Fig. 1.34.

135

2. PRINCIPAL ACCOMPLISHMENTS IN THE PAST YEAR

Most of the accomplishments of the past year will be described in detail in succeeding sections. They are listed here in concise form. Activities during the present year have been grouped under five major areas - Thermal Oxidation, Ion Implantation and Diffusion, Chemical Vapor Deposition, Materials Research and Interface Physics, and SUPREM Implementation.

2.1 THERMAL OXIDATION

1. A study of the high pressure oxidation kinetics of silicon in both pyrogenic and dry O_2 ambients has been completed. Linear pressure dependence for both B and B/A in the linear-parabolic model was found for H_2O oxidation. In the case of dry O_2 oxidation, B was found to be linearly proportional to pressure, while B/A showed a partial pressure dependence of the form p^n where $0.7 < n < 0.8$. This data has been modeled using the linear parabolic growth law and incorporated in SUPREM-III.

2. A large body of kinetic data covering the thin oxide ($x_0 < 500$ A in dry O_2) regime has been gathered in collaboration with E. Irene of IBM. These experiments made use of the IBM in-situ ellipsometer system and covered the temperature range 800 - 1000°C, and partial O_2 pressures of 0.01 - 1.0 atm. This data is still being analyzed, but it appears that it can be described to first order by a simple decaying exponential kinetic law. This model has been incorporated in SUPREM-III.

3. Experimental measurements of fixed oxide charge density versus growth temperature for as-grown and Ar annealed samples have been carried out over the temperature range 700 - 1200°C. The data is not sufficiently complete as yet to immediately incorporate a model in SUPREM III, but the

results are sufficiently encouraging that we expect such a model will be forthcoming in the future. Q_f shows both a growth rate dependence $\alpha(dx/dt)^n$ similar to OED and OISF data, and an exponential dependence on temperature.

4. A theoretical analysis of the thermal oxidation process has indicated that many of the apparently contradictory results in the literature regarding molecular vs. atomic oxidant species and the effects of electric fields on the process, can be resolved if the oxidation process is modeled as occurring via a two layer SiO_2 structure. The bulk of the SiO_2 is stoichiometric and amorphous. A thin layer near the Si/SiO_2 interface is non-stoichiometric (Si rich) and acts as a dense blocking layer for the oxidant diffusion process.

2.2 ION IMPLANTATION AND DIFFUSION

1. The Boltzmann Transport equation (TE) modeling of ion implanted range statistics for ion energies >10 KeV has been completed. A simplified version has been adapted for and incorporated into SUPREM-III.

2. TE Modeling of the range statistics of knock-on ions (recoils) and channeling tails has shown excellent agreement with experiment in the first case and promising results in the second, although further work is required. In the specific case of oxygen knock-ons from an SiO_2 mask, simplified expressions suitable for SUPREM have been developed although they will not be implemented in the first released version of SUPREM-III. We expect to incorporate a general knock-on modeling capability in future updated versions of the program as more complete data becomes available.

3. Models for intrinsic and extrinsic (i.e., doping dependent) diffusion coefficients for B, P, As and Sb have been improved and incorporated in

SUPREM-III. This largely involved a careful literature study to sort out the best current thinking on diffusion in silicon.

4. Models for oxidation enhanced diffusion (OED) for B, P and As and oxidation retarded diffusion (ORD) for Sb have been implemented in SUPREM-III. These models are based partly on experimental work done at Stanford, but they also incorporate a wide body of data from the literature.

5. A temperature tolerant metallurgy using WSi_2 has been developed for experimental characterization of transient process phenomena. Both MOS and PN diode structures have been found to be stable even after repeated temperature cycling up to 1000°C. This technique is currently being applied to a study of backside gettering (and the roles interstitials and vacancies play in the process) and to a study of oxide charge annealing.

2.3 CHEMICAL VAPOR DEPOSITION

1. Models for dopant segregation at polysilicon grain boundaries, carrier trapping at grain boundaries, and grain growth have been combined to develop an overall predictive capability for resistivity of polysilicon thin films. A simplified version of this model has been implemented in SUPREM-III.

2. Work in progress has already established a qualitative understanding of the oxidation kinetics of polysilicon thin films. A series of experiments has demonstrated that for both lightly and heavily doped polysilicon films oxidation kinetics can be modeled as in single crystal substrates, with the electrically active dopant concentration determined from the resistivity model described above. The electrically active dopant concentration determines the total vacancy concentration and hence B/A in heavily doped films. This model has been implemented in SUPREM-III.

138

3. Diffusion from doped polysilicon sources into single crystal substrates (i.e., buried contacts) has been experimentally studied. It has been found that OED effects when the top polysilicon surface is oxidized are similar to single crystal structures, except that an additional term must be included in the model to account for interstitial recombination at polysilicon grain boundaries.

4. Preliminary experimental results to understand $MoSi_2$ and $TaSi_2$ oxidation kinetics and silicide/single crystal contact resistance and dopant segregation, have been obtained. The results, however, are too preliminary to include models for these structures in the first versions of SUPREM-III.

2.4 MATERIALS RESEARCH AND INTERFACE PHYSICS

1. AES and SIMS measurements of the Chlorine distribution in SiO_2 during O_2/HCl oxidation have been completed. It has been found that the Chlorine piles up in a region < 30 Å from the Si/SIO_2 interface, in the SiO_2. The profile in the SiO_2 has been measured using SIMS and a simple analytic model developed to explain the results. The measurements point to the presence of a field driven transport process.

2. A comparison of AES, SIMS, RBS and XPS has been completed. The relative advantages of each technique have been tabulated enabling a specific technique to be chosen as optimum for a particular interface or analysis problem.

3. An empirical model for P pile-up near the Si/SiO_2 interface (in the silicon) has been developed based upon a large body of experimental measurements made under this contract during the past two years. This model has been incorporated in SUPREM-III and essentially treats the interface as a large sink for P and other N-type dopants.

139

2.5 SUPREM IMPLEMENTATION

1. The overall highlight here has been the re-writing of the SUPREM code to model multi-layer structures. Substantially improved and in many cases completely new models for diffusion, oxidation, OED, ion implantation, Si_3N_4 and polysilicon have been implemented in SUPREM-III.

2. The models used in SUPREM for calculation of electrical parameters (MOS threshold voltage, ρs, C, etc.) have been greatly improved over SUPREM-II primarily by incorporating a rigorous Poisson solution in the program. These changes have been implemented in SUPREM-III.

3. Both the SEDAN (1D) and GEMINI (2D) device analysis programs have been released and widely distributed. Both are directly compatible with SUPREM outputs.

DETAILED

TECHNICAL

REPORTS

3. HIGH PRESSURE OXIDATION OF SINGLE-CRYSTAL SILICON IN DRY O_2

L. Lie, R. Razouk, B. Deal

3.1 INTRODUCTION

The increasing demand for improved VLSI device performance and higher packing density is leading towards an increased reliance on low temperature-shorter time processing. High pressure oxidation has been gaining wider acceptance in the last year as a standard tool in the fabrication of integrated circuits. Aside from the oxidation time reduction associated with high pressure oxidation, the substantial reduction in dopant diffusion and redistribution makes this process particularly suitable for application in the fabrication of structures with submicron features in both MOS and bipolar devices. Furthermore, results published in the literature by Tsubouchi et al., [3.1] and by Katz and Kimerling [3.2] indicate reduction in oxidation-induced stacking faults as a result of the lower temperatures possible with high pressure oxidation. Finally, the successful application of high pressure oxidation of silicon to the fabrication of MOS devices in steam as well as in dry O_2 ambients has been realized [3.3,3.4].

In the last two years, we have investigated the high pressure oxidation of single crystal <100> and <111> silicon in both dry O_2 and pyrogenic H_2O ambients [3.5,3.6]. A linear pressure dependence was found for both the parabolic rate constant B and the linear rate constant B/A in the case of silicon oxidation in pyrogenic H_2O as shown in Fig. 3.1. In the case of dry O_2 oxidation, B was found to be linearly proportional to pressure, while B/A showed a pressure dependence of the form P^n where $0.7 < n < 0.8$ as indicated in Fig. 3.2. The results of this investigation are incorporated in SUPREM III and were described in detail in last year's annual report. In this report, we briefly examine

examine the implication of the pressure dependence observed on the physical variables of the thermal oxidation mechanism.

3.2 DISCUSSION AND CONCLUSIONS

Oxidations of silicon in dry O_2, although not as rapid as in steam, can result in improved oxide properties in terms of reduced electron trapping [3.8] and silicon defects [3.9,3.10]. Mobile ionic charges may also be reduced with dry O_2 oxidations. With the use of high pressure, thereby, it is possible to grow high quality dry O_2 oxides at rates comparable to steam oxidations at the same temperature.

As reported earlier [3.5] high pressure oxidation of lightly doped <100> and <111> silicon in dry O_2 have been investigated over the temperature range of 800° to 1000°C and pressure range of 1 to 20 atm. The power dependence of the linear rate constant B/A was found to be markedly different from that of silicon oxidation in pyrogenic H_2O over the same temperature and pressure range as can be seen in Figs. 3.1 and 3.2. This result which can be represented as

$$B \propto P^n \quad , \quad n = 1 \tag{3.1}$$

$$B/A \propto P^m, \quad m = 0.7 \tag{3.2}$$

and is related to the physical variables involved in the oxidation process as follows [3.7]: $B = 2 DC^*/N_1$ and $B/A = C^*/(N_1(1/k + 1/h))$, where B and B/A are the parabolic and linear rate constant, respectively, D is the effective diffusivity of the oxidant in the oxide, C^* is the equilibrium concentration of the oxidant in the oxide, N_1 is the number of oxidant molecules incorporated into a unit volume of the oxide layer, k is a first-order chemical surface reaction constant, and h is the gas-phase transport coefficient. In terms of the proportionality of these variables to the oxidant pressure, eqs. (3.1) and (3.2) can then be written as:

$$B \propto DC^* \propto P \tag{3.3}$$

$$B/A \propto C^*k \propto P^{0.7} \tag{3.4}$$

Thus $\quad A \propto Dk^{-1} \propto P^{0.3}$ \hfill (3.5)

where the assumption is made in writing equations (3.4) and (3.5) that the gas phase transport coefficient (h) used to describe the flux of oxidant at the oxide-bas interface is sufficiently large to have a negligible effect on the linear oxidation rate constant [3.7].

It is of interest to examine the physical variables outlined in equations (3.4) and (3.5) and to attempt to relate these variables to the pressure dependence obtained experimentally. In the pressure range examined, which can be considered quite moderate, the diffusion constant D for the oxidant in the solid phase can be assumed constant to a first order approximation. The pressure dependence of B would therefore come exclusively through C* as outlined in the Deal-Grove linear parabolic model. This implies that the equilibrium concentration of the oxidant in the oxide obeys Henry's law in that pressure range. This conclusion is further strengthened by results obtained in the steam oxidation of single crystal silicon [3.5] and doped polycrystalline silicon (described in other sections of this report) which have indicated that the parabolic rate constant is linearly proportional to pressure in the range of 1 - 25 atm. The dependence of the equilibrium oxidant concentration on pressure as outlined above also implies that based on the Deal-Grove linear parabolic model, the oxidant transport through the oxide involves primarily molecular species rather than dissociate ionic species. It should be emphasized that the analysis carried out here is based entirely on the linear-parabolic Deal-Grove oxidation model and we will not address at this stage various oxidation mechanisms which have been proposed in the literature and which deal with a variety of diffusing species and electrical or electrochemical fields present during

the oxidation and which may lead to the varying pressure dependences of the oxidation rate constants.

Considering the pressure dependence of C* and D discussed above, equations (3.4) and (3.5) suggest that the first order chemical surface reaction constant (k) is proportional to pressure to a power of about (-0.3). This implies that although the oxidation rate constant (B/A) and the oxidation rate both increase with increasing pressure, this increase is primarily due to the increase in oxidant concentration and that the effective chemical surface reaction constant is actually reduced at increased pressures over 1 atm. Such a pressure dependence would imply a complex interface reaction mechanism coupled with a possible saturation in the interfacial flux across the Si/SiO_2 interface. This is in contrast to our analysis of the oxidation of silicon in steam in the same pressure range [3.5]. In these experiments, using the same analysis carried out here, we found the chemical reaction rate constant invariant with pressure. The difference between the steam and dry oxygen oxidation mechanism hinges on differences in the oxidizing species as well as the presence of hydrogen as a subproduct component in the overall reaction in steam. The difference in oxidation mechanism is also reflected in the initial rapid oxidation regime observed in dry O_2 oxidation and which is not present in most steam oxidations with the possible exception of lower oxidation temperatures.

Other arguments for the 0.7 power dependence of the linear rate constant may come from consideration of the interfacial flux at the SiO_2/gas interface. As pointed out previously, equation (3.4) assumes a negligible contribution due to gas phase transport to the linear oxidation rate constant (B/A). Should that condition change with increasing pressure above 1 atm the effect could be to reduce the effect linear rate constant.

Finally, it should be pointed out that the effect of oxidant pressure on the kinetics of silicon oxidation in dry O_2 as determined from this work and

summarized in expressions (3.1) and (3.2) was also found to model with an accuracy of better than 3% the oxidation data at 130 atm obtained by Zeto et. al. [3.13].

In summary, high pessure oxidation of silicon in a dry O_2 ambient over the temperature range of 800° to 1000°C and the pressure range of 1 to 20 atm has been characterized. These data have been modeled using the linear-parabolic growth law and are incorporated in SUPREM III. The implications of the rate constant power dependences on the variation of oxygen diffusivity and solubility in the oxide with pressure have been discussed.

REFERENCES

[3.1] N. Tsubouchi, H. Miyoshi, and H. Abe, Jap. J. Appl. Phys., 17, Suppl. 17-1, 223 (1978).

[3.2] L. E. Katz and L. C. Kimerling, J. Electrochem. Soc., 125, 1680 (1978).

[3.3] N. Tsubouchi, H. Miyoshi, H. Abe, and T. Enomoto, IEEE Trans. Electron Devices, ED-26, 618 (1979).

[3.4] S. Marshall, R. J. Zeto, and J. S. Kesperis, Electrochemical Society Spring Meeting, Abstract No. 83, 1977.

[3.5] J. D. Plummer et. al., "Final Report on Computer-Aided Design of Integrated Circuit Fabrication Processes for VLSI Devices," Stanford Electronics Labs, TR DXG501-81, July 1981.

[3.6] R. R. Razouk, L. N. Lie, and B. E. Deal, J. Electrochem. Soc., 128, 2214 (1981).

[3.7] B. E. Deal and A. S. Grove, J. Appl. Phys., 36, 3770 (1965).

[3.8] R. G. Gdula, J. Electrochem. Soc., 123, 41 (1976).

[3.9] H. J. Queisser and P. G. G. van Loon, J. Appl. Phys., 35, 3066 (1964).

[3.10] A. Mayer, RCA Review, 31, 414 (1970).

[3.11] R. J. Zeto, E. Hryckowian, C. D. Bosco, G. J. Iafrate, R. W. Brower, and C. G. Thornton, Electrochemical Society Fall Meeting, Abstract No. 206, 1974.
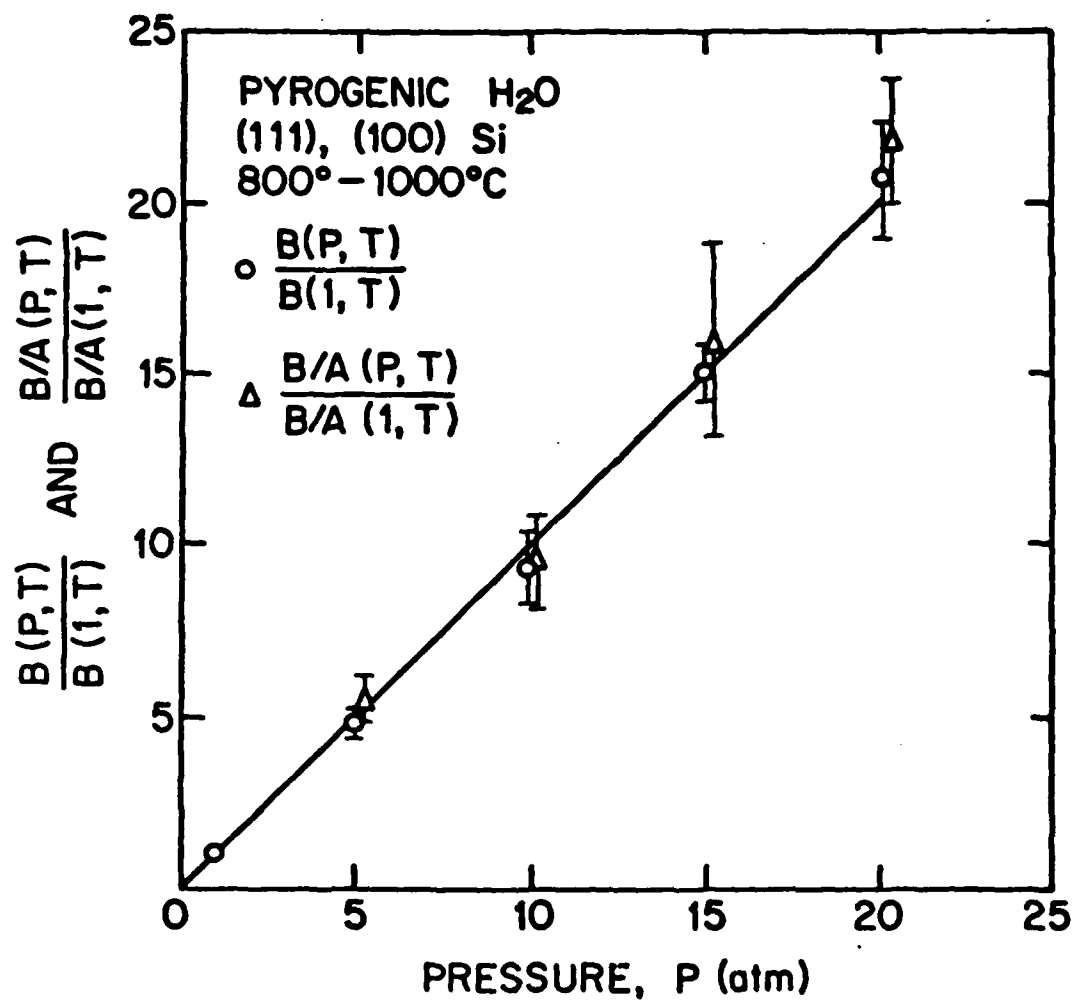
Fig. 3.1: Normalized parabolic and linear rate constant B(P,T)/Bl,T) and [B/A(P,TP]/[B/A(1,T)] vs H$_2$O pressure for <100> and <111> silicon oxidation in pyrogenic H$_2$O ambient at 1, 5, 10, 15, and 20 atm. Temperature range was 800°C - 1000°C.
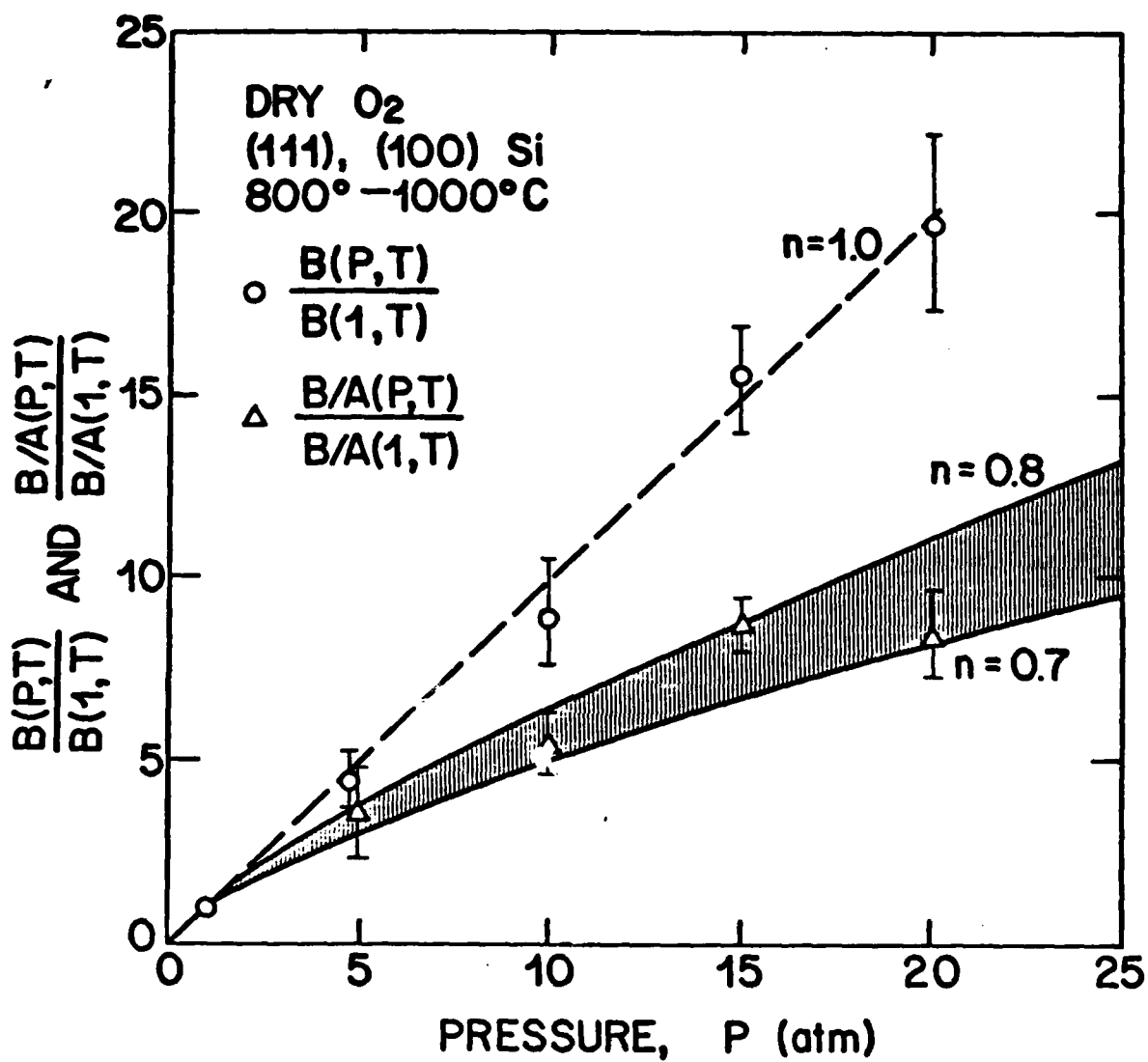
Fig. 3.2: Normalized parabolic and linear rate constant $B(P,T)/B(1,T)$ and $[B/A(P,T)]/[B/A(1,T)]$ vs O_2 pressure for <100> and <111> silicon oxidized in dry O_2 at 1, 5, 10, 15, and 20 atm. Temperature range was 800°C - 1000°C.

4. OXIDATION OF DOPED POLYCRYSTALLINE SILICON

R. Razouk, K. Saraswat, L. Lie, E. Kutlu, B. Deal, J. Plummer

4.1 INTRODUCTION

The trend toward the process modeling of multilayer structures as well as two-dimensional structures is becoming quite evident as we advance to more complex VLSI structures. In a typical VLSI structure, one or more polycrystalline silicon films can be used as interconnect between various devices (on one or more levels) and as parts of active devices such as resistors and transistors. Passivation between the layers is typically accomplished by oxidation of the poly-silicon film itself or by using the poly-silicon as a source of silicon in the thermal oxidation of various metal silicides. The characterization of the kinetics of oxidation of polycrystalline silicon films and the electrical properties of both the films and their oxides are therefore quite important to the continued advancement of VLSI circuit fabrication technology.

The thermal oxidation of single-crystal silicon has been extensively investigated over the last two decades. The Deal-Grove linear-parabolic growth law has been the most accepted kinetic model for single crystal silicon oxidation. The oxidation behavior of polycrystalline silicon, however, differs from that of single crystal silicon oxidation as a result of the small crystallite formation in the polycrystalline silicon structure. Grain growth, recrystallization, dopant segregation and diffusion, carrier trapping at the grain boundaries, carrier mobility, and density of electrically active dopant are additional factors to be taken into account in the oxidation mechanism of poly-Si. Variations in the overall physical and electrical properties of polycrystalline silicon resulting from doping and/or thermal anneal, and oxidation can be related to these factors. Continuing effort has been made in many laboratories to understand the interactive correlations among these factors, principally the effects of dopant carriers

on the film resistivity for which two main models have been proposed to date, namely the carrier trapping [4.1-4.4] and the dopant segregation models [4.5-4.7].

Many publications on atmospheric oxidation of poly-Si in the last few years [4.8-4.12] illustrate the general interest and need in the investigation of poly-Si oxidation. The influence of the initial poly-Si film thickness on the oxidation behavior of poly-Si films has also been observed [4.13,4.14]. A thorough characterization of the effect of high pressure on the oxidations of poly-Si, however, is still needed.

Work on the characterization of the effect of pressure on the oxidation kinetics of doped and undoped polycrystalline silicon films is the objective of the first part of this chapter. The oxidation of phosphorus doped LPCVD polycrystalline silicon films at pressures in the range of 1 to 10 atm pyrogenic H_2O over the temperature range of 750° to 850°C have been characterized to date and is described below. The second part of this chapter describes the efforts to model the kinetics of the oxidation of polycrystalline silicon.

4.2   HIGH PRESSURE OXIDATION OF DOPED POLYCRYSTALLINE SILICON IN PYROGENIC H_2O

4.2.1   Experimental

The starting material for the preparation of the polycrsytalline silicon films used in this work was four-inch diameter p-<100> silicon wafers of 5-9 Ω-cm resistivity. The wafers were oxidized in a pyrogenic H_2O/5% HCl ambient at 920°C to grow 1060 A oxides, followed by the deposition of LPCVD poly-silicon films at 625°C with thicknesses ranging from 5800 to 6050 A. Half of the total of 144 wafers were doped in PH_3 at 950°C for 40 minutes followed by a 10 minute nitrogen purge. After removal of phosphosilicate glass, poly-silicon film thickness was generally reduced by 500 to 600 A. Half of the doped and half of the undoped samples were annealed in a non-oxidizing ambient in order to stabilize the grain size. Prior to annealing, 0.32 µm vapox caps were deposited at 450°C on these

wafers to protect the surface of the film and to prevent dopant outdiffusion. Pre-oxidation thermal anneals were carried out in N_2 at 1000°C, 900°C, 850°C, 800°C, and 750°C for 2, 13, 24, 64, and 140 hr, respectively [4.7]. The vapox protective caps were then removed in a buffered SiO_2 etch following the anneal.

Resistivity measurements were performed on the polysilicon films by means of a four point probe. Doped and annealed films showed variations from 5.0 to 6.3 x 10^{-4} Ω-cm with the 1000° anneal yielding the lowest resistivity and the 750°C anneal the highest resistivity, as shown in Fig. 4.1. Doped and unannealed samples showed variations from 5.0 to 5.5 x 10^{-4} Ω-cm. Typical standard deviation of the resistivity measurements was ±3%.

Polysilicon film thickness was measured by a NanoSpec spectrophotometer. The average thickness of the films were measured at 5980 Å for undoped unannealed samples, 5700 Å for undoped annealed, and 5450 Å for doped unannealed as well as annealed. The thinner results obtained for the doped films were due to the phosphosilicate glass removal. Thickness uniformity was generally better than ±2%.

A Gasonics high pressure oxidation system which has been previously described [4.13,4.15] was used to generate the polysilicon oxide growth data. The same standard wafer cleaning and drying procedures, which consisted of submersion in solutions of H_2O/H_2SO_4, aqua regia, and 10:1 H_2O/HF with appropriate DI water rinses following by spin-dry in nitrogen, were carried out prior to oxidations. Figure 4.2 shows a typical high pressure oxidation cycle used to oxidize the phosphorus doped poly-silicon in pyrogenic H_2O. Argon was used to pressurize the system, instead of the oxidizing pyrogenic H_2O ambient as was done in the previous work dealing with the characterization of high pressure oxidation of lightly doped single crystal silicon in pyrogenic H_2O. Since the heavily doped thin polysilicon films are of 5500 Å average thickness, the polysilicon oxide thicknesses grown were limited to less than about 8000 Å. Due to the high oxidation rate of doped polysilicon in steam at high pressure and in order to ensure

151

maximum accuracy, system pressurization was carried out in argon. Once the tube pressure had stabilized at the required pressure, steam was introduced. It should be noted that a time Δt_1, estimated to be less than 5 minutes, elapses before the ambient in the tube is completely changed from argon to steam as shown in Fig. 4.2. At the end of the oxidation cycle, the ambient was switched back to argon at high pressure for 5 minutes prior to depressurization. In this manner the effect of the ambient exchange time Δt_1, on the oxidation kinetics is minimized. A lightly doped p-<100> single crystal silicon wafer was included in every oxidation run as a monitor. Oxide thickness grown on the polysilicon films was measured by an Alpha-Step profilometer.

4.2.2 Results and Discussion

4.2.2.A Oxide Thickness Data

Oxidation data of phosphorus doped poly-Si films at 1, 5, and 10 atm pyrogenic H_2O and at temperatures of 750°, 800°, and 850°C are shown in Figs. 4.3, 4.4, and 4.5. The two sets of curves at each temperature represent the oxidation data of the poly-Si samples which received thermal anneal in N_2 prior to oxidation and those which received no anneal. The thickness of oxides grown on the poly-Si films ranged from 300 Å to 6000 Å. The results show consistently lower oxidation rates for the samples which received a pre-oxidation anneal. This might be related to the effect of dopant segregation at the grain boundaries, which reduces the electrically active dopant concentration in the grains, as a result of the pre-oxidation anneal [4.7]. The reduction of the electrically active dopant concentration in the grains is reflected in the higher resistivity of the poly-Si samples annealed at 750°, 800° and 850°C as observed in Fig. 4.1. The effect of the electrically active doping level on increasing the oxide growth rates of single crystal silicon has been observed and comprehensively characterized by Ho and Plummer [4.16,4.17]. In addition to the effect of dopant

segregation, the difference in oxidation rates between the two poly-Si films might be also due to a difference in grain size.

Comparison between the data obtained at Fairchild at 1 atm and that obtained by K. Saraswat of Stanford and reported last year in this program [4.13], indicates substantial variations at all three temperatures in the oxidation rates for polySi films of similar initial thickness and resistivity. Figure 4.5 shows comparison of the data at 800°C. The oxidation rates of the 0.55 μm initial polysilicon film thickness are lower than both the 0.5 μm and the 1.0 μm initial polysilicon film thickness. Investigation of the cause of the variation is being carried out. Structural evaluation of the polysilicon films using TEM and X-ray diffraction measurements should help in determining the effect of grain size and orientation, if any, on the oxidation of polysilicon. Differences in grain size and orientation could result from different processing parameters such as source, temperature, and length of doping and annealing.

4.2.2.B Oxidation Rate Constants

The determination of the oxidation rate constants was carried out using the Deal-Grove linear-parabolic model [4.18] which can be written as

$$x_o^2 + Ax_o = B t + x_i^2 + Ax_i \tag{4.1}$$

with $t = t - t_i$, $x_i = 0$, and $t_i = 0$ for oxidations in pyrogenic H_2O. $B = B(T,P)$ includes the temperature and pressure factors of the oxidation. The plot of log parabolic rate constant B versus reciprocal oxidation temperature for phoshporus doped poly-Si oxidations in pyrogenic H_2O for all pressure and temperature ranges investigated is shown in Fig. 4.7. The parabolic rate constants B for doped poly-Si are higher than that of lightly doped single crystal silicon, especially at lower temperatures. For the same temperature range (T < 900°C), the activation energy of B for H_2O oxidations of doped poly-Si is 16 kcal/mole

(0.69 eV) which is lower than that of lightly doped <100> and <111> single crystal silicon (26 kcal/mole or 1.1 eV). The activation energy of B in the lower temperature range of the doped poly-Si oxidation approaches that of the higher temperature B of lightly doped single crystal silicon. A decrease in activation energy at lower temperatures for heavily doped single crystal silicon oxidized in a dry O_2 ambient has been previously observed by Ho and Plummer [4.19]. It was also then noted that this decrease in activation energy increases with increasing doping level. Thus, the decrease in activation energy found for the oxidation of the doped poly-Si might be related to the effect of high doping concentration in the substrate, and the segregation of these dopants into the SiO_2.

The graph of log linear rate constant B/A plotted versus reciprocal oxidation temperature is shown in Fig. 4.8. Higher values of B/A compared to those of the lightly doped <100> and <111> single crystal silicon at a given temperature and pressure are also noted. The activation energy for B/A of doped poly-Si is 41 kcal/mole (1.8 eV). This is similar to that of lightly doped <100> and <111> dingle crystal silicon (45 kcal/mole or 1.9 eV). This similarity would be expected since the activation energy of B/A should not vary with grain orientation. It was also noted that up to 10 atm the activation energies of B and B/A of doped poly-Si in pyrogenic H_2O are found to be independent on the oxidation pressure.

The graphs of normalized B and B/A plotted versus oxidation pressure are shown in Figs. 4.9 and 4.10, respectively. For the entire range of oxidation parameters investigated, B was found to be linearly proportional to pressure within the accuracy of the measurements. The same relationship was found for oxidations of lightly doped single crystal silicon either in pyrogenic H_2O or dry O_2 pressures up to 20 atm in the temperature range of 800° to 1000°C. The pressure dependence of B/A, however, indicates proportionality to pressure to the

power of n with 0.7 < n < 0.8. This relationship deviates from that of oxidation of lightly doped single crystal silicon in pyrogenic H_2O where a linear dependence of B/A on H_2O pressure was observed. On the other hand, it is interesting to note that an identical less-than-linear relationship for B/A to pressure occurs in the oxidation of lightly doped single crystal silicon in dry O_2 [4.13,4.20]. Results from the next sets of data on the high pressure oxidation of undoped polycrystalline silicon in pyrogenic H_2O and doped and undoped polycrystalline silicon in dry O_2 will undoubtedly enable further analysis and lead us to a better understanding of the kinetics and the effects of high pressure on the oxidations of polycrystalline silicon films.

4.3 OXIDATION OF DOPED POLYCRYSTALLINE SILICON IN DRY O_2

Last year we reported experimental results on the thermal oxidation of heavily phosphorus doped polycrystalline silicon films ·n dry oxygen [4.13,4.14]. In this report a discussion of the kinetics of oxidation is presented. The results reported were for polycrystalline films with thicknesses of 0.5 and 1.0 μm which were deposited onto thermally oxidized silicon wafers by low pressure chemical vapor deposition (LPCVD) at 620°C and were doped to solid solubility with phosphorus by diffusion from a $POCl_3$ source at 1000°C for 30 min. Oxidations in dry O_2 were carried out at temperatures of 1100°, 1000°, 900° and 800°C for durations of 30 min. to 88 hr. In our earlier work in this program we have shown that dopant segregation at the grain boundaries can reduce the number of electrically active dopant atoms in polysilicon substantially, and the effect is stronger at lower temperatures [4.7]. For the 800°C dry O_2 oxidation, therefore, additional samples were used where a 64 hr anneal in N_2 at 800°C was done prior to the oxidation to segregate phosphorus at the grain boundaries. This duration was sufficient to complete the segregation process [4.7].

For the oxide thickness measurements of polysilicon samples, steps were first etched in SiO_2 and then the thickness was measured with an Alphastep pro-

filometer. The thickness of SiO_2 on single-crystal control wafers was measured with both an ellipsometer, and the Alphastep profilometer. Agreement within 5% was obtained between the two methods.

The results of oxidations in dry oxygen are reproduced in Figs. 4.11 and 4.12. It is obvious from Fig. 4.11 that heavy doping enhances the rate of oxidation of polysilicon relative to undoped single-crystal silicon and the enhancement is stronger at lower oxidation temperatures. This behavior is similar to that observed in doped single-crystal silicon; however, if the data of Figs. 4.11 and 4.12 are compared to those for the oxidzation of phoshporus-doped single-crystal silicon [4.19], it can be concluded that the enhancement of oxidation resulting from phoshporus doping is much greater for polysilicon than it is for single-crystal silicon. This is more obvious at lower oxidation temperatures. As the temperature is increased, the differences between doped polysilicon and doped and undoped single crystal silicon become smaller. This can be seen very clearly for the case of 1100°C dry oxidation.

Very little data on similarly doped polysilicon are available in the literature for direct comparison to our data. Sunami's work [4.8] was focussed on the wet oxidation (O_2 bubbled through 90°C water) of ion implanted polysilicon. Kamins [4.9] has only one data point, 850°C steam oxidation of heavily doped polysilicon for 75 min, and it is in good agreement with our results. The results obtained by Barnes et. al., [4.10] for 800°C dry O_2 oxidation of phosporus-doped polysilicon are not in complete agreement with our work. The oxidation rate was slightly lower for ther films doped at 950°C as compared to 1000°C in our work, which may have resulted in a lower total dopant concentration, different grain structure and size, and therefore, a slower oxidation rate.

An approximate relationship $X_0 \approx Kt^n$ has been shown by Sunami [4.8] to model the oxidation of polysilicon, where K and n are empirical constants used

156

for curve fitting. For undoped silicon, n is always between 1.0 (linear oxidation) and 0.5 (parabolic oxidation). It is evident in Figs. 4.11 and 4.12 that the slope of the X_0 vs t curves for polysilicon is less than that for single crystal silicon, i.e., n < 0.5, and similar lower values of n can be observed in the data of Barnes, et. al. [4.10] for heavily P- and As-doped polysilicon. It can be seen in Fig. 4.12 that n is a function of annealing and thickness of the films. This is purely an empirical approach that merely indicates the differences in the oxidation of polysilicon and single-crystal silicon.

Values of oxidation rate constants A, B and X_i were calculated using the conventional procedure [4.18]. X_i was calculated by plotting the oxide thickness against oxidation time as shown in Fig. 4.13 for the case of 800°C oxidations. Table 4.2 summarizes the values of X_i. It is interesting to note that X_i is dependent upon polysilicon thickness; the thicker the polysilicon is the lower the value of X_i becomes. Furthermore, the X_i values are much higher than for single crystal silicon [4.18,4.19]. These results imply that during the oxidation process, the initial oxidation regime may be dominant for a longer period of time for polysilicon as compared to single crystal silicon. From the plots of $(X_0 + X_i)$ vs. $t/(X_0 - X_i)$, values of A and B were calculated. One example is shown in Fig. 4.14. It was found that the values of A were close to zero and sometimes even negative and highly dependent on film thickess. Similarly, X_i was a function of film thickness and temperature. Small or negative values of A have no physical meaning. We might assume infinite linear rate constant. The values of B were fairly consistent with published results in the literature and are listed in Table 4.1; comparison data from the literature are also included. It can be seen that the values of B are significantly higher compared to undoped single-crystal silicon but only slightly higher compared to doped silicon. This difference is more noticeable at lower temperatures. The increase in

B with doping is consistent with those observed by Ho et. al. [4.19] for the oxidation of heavily phosphorus-doped single-crystal silicon. Calculated parabolic rate constants are plotted as a function of temperature in comparison with the doped and undoped single crystal parabolic rate constants in Fig. 4.15 for dry oxidation. Although the data points are scattered around the single crystal B values, there is nothing particular about this scattering. So, it can be concluded that the parabolic rate constants of polysilicon and single crystal silicon are almost the same, which implies that the diffusion of the oxygen molecules through the oxide is roughly the same for the oxidaton of the two materials.

Oxide thicknesses are calculated as a function of time by using the extracted initial oxide thicknesses and rate constants calculated above, and the results are given in Fig. 4.16. It can be seen in Fig. 4.16 that the Deal-Grove model [4.18] can be used to model polysilicon oxidation kinetics for thicknesses higher than 1000 Å provided large values of X_i are assumed and B/A is assumed to be infinite.

Fig. 4.12 plots the effect of the 64-hour preanneal on the oxidation of polysilicon in dry oxygen at 800°C. It can be seen that, for the preannealed polysilicon, the oxides are thinner. This preanneal increased the sheet resistance of the 0.5 and 1.0 μm thick films from 10.6 and 4.8 ohms/sq to 12.7 and 5.3 ohms/sq, respectively, which can be attributed to dopant segregation at the grain boundaries [4.7] and a reduction in solid solubility [4.21]. These two phenomena will be discussed in more detail later on in this section. These two phenomena occur gradually as the oxidation proceeds because they are very slow processes and require extremely long times (especially segregation) to reach saturation [4.7,4.22,4.23]. Because attempts to obtain variations in the value of B were not succssful because of scatter in the data, it cannot be ascertained how B is affected by dopant segregation to the grain boundaries caused by the pre-

158

anneal. Large variations in A and X_j were observed; however, meaningful values could not be calculated for reasons elaborated earlier.

4.4 MECHANISM OF OXIDATION OF DOPED POLYCRYSTALLINE SILICON

In the final report of last year [4.13] the sheet resistance and resistivity of the remaining layers of polysilicon after oxidation were reported and from the results, the following observations were made:

1. The electrically active dopant concentration in polycrystalline silicon is affected by processing subsequent to the doping because of dopant segregation to the grain boundaries [4.7,4.22] and changes in the solid solubility [4.21,4.23)]. These changes affect the kinetics of oxidation.

2. The grain boundaries can be oxidized preferentially leading to an enhanced average oxidation rate of polysilicon [4.12].

The major observations not fully explained are that (1) oxides grown on 0.5 μm thick films are consistently thicker than those grown on 1.0 μm thick films and (2) polysilicon can be oxidized faster than similarly doped single-crystal silicon. These observations cannot be explained merely on the basis of the electrical activity of the dopant. It is evident from the resistivity measurements that the electrically active doping density of the 0.5 μm thick film is approximately the same as that of the 1.0 μm thick film. If the slightly higher resistivity of the 0.5 μm thick films could be attributed to lower electrically active doping density, it should result in thinner oxides. As pointed out earlier, the higher resistivity has been attributed to lower carrier mobility because of the smaller grain size [4.24]. Baldi et. al. [4.11] in their work on oxidation of arsenic doped polysilicon showed that the oxidation rate was significantly influenced by the chemical concentration of arsenic, which was in the range of 1.2 to 4.7 x 10^{21} cm^{-3}. At these heavy doping levels, the electrically active

arsenic concentration is expected to be limited by both the electrical, as well as the total solubility [4.21,4.23]. Therefore, their observations cannot be explained merely on the basis of an enhanced oxidation resulting from an increase in vacancies caused by a shift in the Fermi level because, for the Fermi level to shift, the density of the electrically active dopant must change. The only other way these and our results can be explained is by assuming that grain boundaries and the dopant present there in unusually large concentrations play a major role in the oxidation process of doped polysilicon.

During the doping of the films, dopant will segregate and precipitate in the grain boundaries, and further processing at lower temperatures will enhance these mechanisms [4.7]. A major fraction of the total dopant, as much as 80%, can segregate to the grain boundaries [4.7]. Mandurah et. al. reported [4.25] that the average width of the grain boundaries is in the range of 7 to 10 Å. In our samples, the grain size is expected to be about 1.0 μm [4.25]. Based on these data, the dopant concentration in the grain bounadries could be in excess of 10^{22} cm^{-3} for an average concentration in the range of 10^{20} cm^3 and this unusually large concentration can enhance the reaction at the grain boundaries between the oxidant and silicon.

In addition to influencing the surface reaction at the grain boundaries, the dopant present there could also affect the oxidant transport through the previously grown SiO_2. It has been demonstrated in several investigations [4.10, 4.11,4.19] that for low and medium dopant concentrations, the parabolic rate constant B is not a function of the dopant concentration; however, with heavy doping it increases with the dopant concentration. This behavior has been attributed to the influence of the dopant in the SiO_2 on the diffusivity of the oxidant [4.19].During oxidation, a fraction of the dopant segregates from Si to the SiO_2. At lower dopant concentrations in Si the concentration in SiO_2 are

not sufficient to influence the diffusivity of the oxidant; however, at higher concentrations in Si, the segregated dopant in the grown SiO_2 is sufficient to increase the diffusivity of the oxidant. This is reflected in an increase in B because it is directly proportional to the oxidant diffusivity [4.16,4.18] and, as is evident in Eq. (4.1), such an increase in B will lead to a thicker oxide. Because, in polysilicon, the dopant concentration in the grain boundaries can be in excess of 10^{22} cm^{-3}, it is probable that a large fraction of this dopant will segregate into the oxide, which would result in an increase in B and therefore, in an increased oxidation rate. Because of the faster oxidation rate at the grain boundaries, the average oxidation rate of polysilicon could be higher than that of similarly doped single-crystal silicon. The conditions that increase the dopant segregation to the grain boundaries will also enlarge this disparity. Mandurah et. al. [4.7] observed that lower temperatures lead to greater segregation. It can be seen in Figs. 4.11 and 4.12 that the enhanced oxidation rate is higher at lower temperatures.

The reason why 0.5 μm thick polysilicon oxidizes faster than 1.0 μm thick polysilicon can be explained. The grain size in the thinner films is always smaller than in the thicker films, and as a result, the number of grain boundaries per unit area at the surface being oxidized is higher in the thinner films. Therefore the undulations in the oxide thickness are closer to each other in the thinner film, resulting in larger average thickness of the oxide. Another factor that may contribute to this behavior is dopant segregation and redistribution at the Si/SiO_2 interface at the grains. During the growth of SiO_2, phosphorus segregates into the grains of polysilicon as it does in single crystal silicon. It has been shown [4.26] that the diffusivity of the dopant is extremely large in the grain boundary and because of this effect any excess dopant in the grains migrates to the grain boundary and quickly redistributes there. The dopant seg-

161

regated at the Si/SiO$_2$ interface probably goes through this process and further modifies the reaction kinetics at the grain boundaries.

The picture that emerges from the preceeding discussion is summarized here. As the oxidation proceeds, the oxidant reaches the SiO$_2$/polysilicon interface, where its reactions with the silicon at the grains and grain boundaries could be different. Because the dopant is electrically active within the grain, the oxidation rate on the surface of the grain should be that of single-crystal silicon. On the other hand, no electrical activity is expected in the grain boundary [4.7, 4.27,4.28]; however, because the structure of the grain boundary differs strongly from that of the grain, its oxidation behavior could also deviate sharply. Dopant concentration in the grain boundaries can be unusually large because of dopant segregation, and this could lead to an enhanced oxidation rate caused by an increase in the surface reaction rate and diffusivity through the grown SiO$_2$. The exact role of the dopant in this process is not clear and further work is required to obtain a better understanding. The rapid diffusion of the dopant atoms in polysilicon also complicates its behavior. The grain boundaries provide a "low impedance path" for diffusion, and this leads to a very quick redistribution of the dopant. All of these effects depend on the grain size, which, in turn, depends on deposition, doping, and high-temperature processing [4.22,4.25]. The grains in the as deposited LPCVD films are on the order of a few hundred Å, however, after doping and/or annealing, their size increases to several thousand Å and the extent of this phenomenon is heavily dependent on the dopant type and density, processing temperature, and time [4.22,4.25]. Therefore, during oxidation the structure of polysilicon will change continuously, and the extent of the effects due to grain boundaries will vary accordingly. Because of this, the models developed for a stable single-crystal silicon structure such as modeled by Eq. (4.1), cannot be applied directly to polysilicon and must be modified. In

162

single-crystal silicon for a given set of oxidation parameters B and B/A can be approximated by constants; however the presence of grain boundaries complicates the case of polysilicon, and B, B/A, and X_i cannot be treated as simple constants. To fully understand the kinetics of polysilicon oxidation, first the role of grain boundaries should be investigated, and only then can a theoretical model be developed.

4.5 MEASUREMENT OF THIN OXIDES ON POLYSILICON

From the preceeding discussion it is evident that B/A and X_i are severely affected by the presence of dopant and the grain boundaries in polysilicon. These two factors influence the linear region of the oxidation process more than the parabolic region. Therefore, in order to fully understand the kinetics of the oxidation process, we must study the growth of thin oxides on polysilicon. From a device point of view, thin oxides are also very important in VLSI circuits.

The measurement of thin oxides on polysilicon is not very straightforward. To begin with, the surface of the polysilicon is very rough and the oxidation process further increases the roughness. In the results reported above, because of the surface roughness of polysilicon, interpretation of the surface profiles generated by the Alphastep profilometer became a tedious task, especially for thinner oxides. The thickness-measurement accuracy of the oxides (on polysilicon) thinner than 1000 A was found to be highly questionable, the variation between various measurements and their interpretation was larger than ± 20%. Therefore, in this work we were forced to avoid very thin oxides on polysilicon. For oxides between 1000 and 2000 A, the variation was less than ± 7.5%, and for greater thicknesses it was negligible.

Furthermore this measurement technique to determine oxide thickness should be reexamined in terms of the thickness undulations observed by Irene et. al.

163

[4.12]. To interpret the data, an average of the varying surface profile is taken, and therefore this technique yields only an average of oxide thickness.

The use of ellipsometry is not possible in this case because of multiple layers confusing the measurement [4.29]. Use of a spectrophotometer is possible only for thicker oxides, because for thinner oxides shorter wavelengths are needed but they are absorbed by silicon and thus no interferen-- pattern can be observed. Thus measurement of thin oxides is not possible by any of the existing conventional techniques. However we must investigate the oxidation kinetics in the thin oxide regime in order to increase our understanding of the mechanisms. Therefore alternative teqhniques to measure the oxide thickness need to be investigated first.

One possible technique is shown in Fig. 4.17 where a capacitor is formed by evaporating aluminum on top of the oxide. By measuring the capacitance, C, between the aluminum and polysilicon electrodes the oxide thicknss, t_{ox}, can be calculated from the relationship

$$C = \epsilon_o K_{ox} \frac{A}{t_{ox}} \qquad (4.2)$$

where ϵ_0 is the permittivity of free space, K_{ox} is the dielectric constant and A is the area of the capacitor.

To determine the feasibility of the technique circular capacitors as shown in Fig. 4.17 were fabricated with 50, 100, 200, 400 and 800 μm radii of the aluminum gates (dots in Fig. 4.17). Initial measurements showed considerable variation in the oxide thickness calculated from Eq. 4.2 as a function of capacitor area. It is believed that the parasitic capacitances and resistance associated with the polysilicon/SiO_2/Si structure enter into the measurements and thus the measured capacitance is different from that predicted by Eq. 4.2. In order to predict the correct value of the Al/SiO_2/polysilicon capacitance, the influence of the parasitic resistance and capacitance must be first evaluated.

The capacitances and the resistances associated with a ring-dot Al/SiO$_2$/poly-Si/SiO$_2$/Si structure are shown in Fig. 4.18. It is evident that besides the Al/SiO$_2$/poly-Si capacitance, C_{ox}, there are many other parasitic elements involved. In particular, the spreading resistance of the polysilicon R_{sp}, oxide capacitance between the polysilicon and the Si substrate C_o, the capacitance of the depletion or inversion layer in the Si substrate C_s and the resistance of the substrate R_s. In general these elements are distributed; however to simplify the analysis we will initially treat them as lumped elements. In the remainder of this section an analysis is presented to help determine the influence of these parasitic elements.

The equivalent circuit of this structure is shown in Fig. 4.19 (a) and the equivalent which is measured is shown in Fig. 4.19 (b), where, C_m and g_m are the measured values of the capacitance and conductance. Obviously C_m can be appreciably different from C_{ox}. Let C_o represent the series combination of C_o and C_s. From the equivalent circuit of Fig. 4.15 it can be deduced that

1. $\lim\limits_{R_{sp} \to 0} C_m = C_{ox}$
$$\tag{4.3}$$

For samples with heavily doped polysilicon, therefore we can expect to measure C_{ox} directly.

2. $\lim\limits_{R_{sp} \to \infty} C_m = \dfrac{C_o' \, C_{ox}}{C_o' + C_{ox}}$
$$\tag{4.4}$$

For samples with lighter doping in the polysilicon the measured capacitance will depend upon C_o' (i.e. C_o and C_s) and therefore on the oxide thickness, t_o, and depletion layer thickness in the Si substrate. In the limiting case where the influence of C_s is assumed to be negligible

165

$$\text{for } t_o \quad t_{ox} \qquad C_m \quad \tfrac{1}{2} C_{ox}$$

$$\text{for } t_x \gg t_{ox} \qquad C_m \quad C_o$$

$$\text{for } t_o \ll t_{ox} \qquad C_m \quad C_{ox}$$

Therefore, to keep the error in the C_{ox} measurement small, we should use thin oxides between the polysilicon and the silicon substrate.

Figs. 4.20 and 4.21 show the results of simulations where C_m/C_{ox} has been calculated as a function of R_{sp}, t_{ox} and t_o for various substrate dopings. By using Figs. 4.20 and 4.21 this technique should yield results for any arbitrary combination of doping density, and substrate oxide thickness t_o.

An easier technique is to measure the parasitics separately. Fig. 4.22 shows two structures fabricated side by side. The structure in Fig. 4.22 (a) is identical to the one in Fig. 4.22 (b) except the oxide on the polysilicon is missing. The basic idea is to measure the equivalent resistance and capacitance of the structure in Fig. 4.22 (a), then use this number in the measurment of the second structure, to calculate the polysilicon capacitance. These techniques are currently being pursued in an investigation of thin oxides on polysilicon.

4.6 SUMMARY AND CONCLUSIONS

High pressure oxidation of phosporus doped polycrystalline silicon in pyrogenic H_2O over the temperature range of 750° to 850°C and pressure range of 1 to 10 atm and atmospheric pressure oxidation in dry O_2 over the temperature range of 800 to 1100°C have been characterized. At atmospheric pressure a variation in oxide growth rates of the poly-Si films originating from two different sources (Stanford and Fairchild) was noted; investigation of the cause of the variation is being carried out at present. For oxidation in H_2O, the parabolic and linear rate constants, B and B/A, were found to be higher than those of lightly doped single crystal <100> and <111> silicon oxidations in pyrogenic H_2O, especially B/A which is higher by approximately an order of magnitude. The activation

166

energy for B was found to be slightly lower compared to that of lightly doped single crystal <100> and <111> silicon oxidations in pyrogenic H_2O; whereas the activation energy for B/A of the doped poly-Si was found to be similar to that of the lightly doped <100> and <111> silicon oxidation. The parabolic rate constant, B, increases linearly with pressure for the conditions investigated in this work. The linear rate constant B/A, however, increases with a less than linear function.

In the case of dry O_2 the rate of oxidation of polysilicon was higher than the correspondingly doped single crystal silicon, values of B were higher than those for lightly doped silicon but comparable to those for heavily doped single crystal silicon and extremely large values of B/A and X_j found.

These results indicate that the thermal oxidation of polysilicon cannot be fully explained by the existing models of oxidation of single crystal silicon. There are two important factors which make the oxidation kinetics of the thin films of polysilicon different--the presence of grain boundaries, and the finite thickness of the film. Because of these two factors, several other mechanisms, such as dopant segregation to the grain boundaries, rapid diffusion of the dopant along the grain boundary etc., influence the kinetics of oxidation significantly, especially in the thin oxide regime.

The use of conventional techniques to measure oxide thickness on polysilicon is limited to about 1000 A. A new technique is being developed where an oxide capacitance measurement yields the oxide thickness.

All of these results taken together will enable a comprehensive and accurate model to be incorporated in SUPREM III for polysilicon oxidation. At the present time, SUPREM III uses essentially the same models for single crystal and poly oxidation. Based on the results of this chapter, it is clear that a better model for polysilicon is needed. The work described here is aimed at such a model.

TABLE 4.1

Parabolic rate constant B (μ^2/min)

Dry O_2 Oxidation

Temp (°C)	This work	Ho et. al. [14] <111> Si @ 8×10^{20} cm^{-3}	Barnes et. al. [10] @ 5×10^{20} cm^{-3}
800	5.1×10^{-5}	4.5×10^{-5}	5×10^{-5}
900	7.2×10^{-5}	8.5×10^{-5}	--
1000	2.34×10^{-4}	1.6×10^{-4}	--
1100	4.5×10^{-4}	4×10^{-4}	--

TABLE 4.2

Oxidation Temp.	X_i (A)	
	0.5 µm Poly-Si	1.0 µm Poly-Si
800°C	1100 Å	600 Å
900°C	600 Å	600 Å
1000°C	1000 Å	500 Å
1100°C	1000 Å	600 Å

REFERENCES

[4.1]   T. I. Kamins, J. App. Phys., 42, 4357 (1971).

[4.2]   J. Y. W. Seto, J. Appl. Phys., 46, 5247 (1975).

[4.3]   G. Baccarani, B. Ricco, and G. Spadini, J. Appl. Phys., 49, 5565 (1978).

[4.4]   G. J. Korsh and R. S. Muller, Solid-State Electron., 21, 1045 (1978).

[4.5]   M. E. Cowher and T. O. Sedgwick, J. Electrochem. Soc., 119, 1565 (1972).

[4.6]   A. L. Fripp and L. H. Slack, J. Electrochem. Soc., 120, 145 (1973).

[4.7]   M. M. Mandurah, K. C. Saraswat, C. R. Helms and T. I. Kamins, J. Appl. Phys., 51, 5755 (1980).

[4.8]   H. Sunami, J. Electrochem. Soc., 125, 892 (1978).

[4.9]   T. I. Kamins, J. Electrochem. Soc., 126, 838 (1979).

[4.10] J. J. Barnes, J. M. DeBlasi, and B. E. Deal, J. Electrochem. Soc., 126, 1779 (1979).

[4.11] L. Baldi, G. Ferla, M. Josi, Electrochemcial Society Spring Meeting, Extended Abstract 168 (1980).

[4.12] E. A. Irene, E. Tierney and D. W. Dong, "Silicon Oxidation Studies: Morphological Aspects of the Oxidation of Polycrystalline Silicon," J. Electrochem. Soc., 127, 705 (1980).

[4.13] J. D. Plummer et. al., "Final Report on Computer-Aided Design of Integrated Circuit Fabrication Porcesses for VLSI Devices," Stanford Electronics Labs, TR DXG501-81, July 1981.

[4.14] K. C. Saraswat and H. Singh, J. Electrochem. Soc., Accepted for publication.

[4.15] R. R. Razouk, L. N. Lie, and B. E. Deal, J. Electrochem. Soc., 128, 2214 (1981).

[4.16] C. P. Ho and J. D. Plummer, J. Electrochem. Soc., 126, 1516 (1979).

[4.17] C. P. Ho and J. D. Plummer, J. Electrochem. Soc., 126, 1523 (1979).

[4.18] B. E. Deal and A. S. Grove, J. Appl. Phys., 36, 3770 (1965).

[4.19] C. P. Ho, J. D. Plummer, J. D. Meindl, and B. E. Deal, J. Electrochem. Soc., 125, 665 (1978).

[4.20] L. N. Lie, R. R. Razouk, and B. E. Deal, J. Electrochem. Soc., submitted for publication.

[4.21] R. B. Fair, "Recent Advances in Implantation and Diffusion Modeling for the Design and Process Control of Bipolar ICs," _Semiconductor Silicon 1977_, Electrochem. Soc., pp. 968-987.

[4.22] M. M. Mandurah, "Physical and Electrical Properties of Polycrystalline-Silicon," Ph.D. Thesis, Stanford Electronics Labs. T.R. No. G503-2, Stanford University, Stanford, CA, June 1981.

[4.23] M. Y. Tsai, F. F. Morehead and J. E. E. Baglin, "Shallow Junctions by High-Dose As Implants in Si: Experiments and Modeling," _J. Appl. Phys._, 51, 3230 (1980).

[4.24] M. M. Mandurah, K. C. Saraswat and T. I. Kamins, "Phosphorus Doping of Low Pressure Chemically Vapor-Deposited Silicom Films, "_J. Electrochem. Soc._, 126, 1019 (1979).

[4.25] Y. Wada and Nishimatsu, "Grain Growth Mechanism of Heavily Phosphorus Implanted Polycrystalline Silicon," _J. Electrochem. Soc._, 125, 1499 (1978).

[4.26] B. Swaminathan, K. C. Saraswat, R. W. Dutton and T. I. Kamins, "Diffusion of Arsenic Polycrystalline Silicon," _Appl. Phys. Lett._, (Submitted).

[4.27] M. M. Mandurah, K. C. Saraswat and T. I. Kamins, "A Model for Conduction in Polycrystalline Silicon, Part I: Theory", _IEEE Trans. Electron. Dev._, ED-33, 1163, Oct 1981.

[4.28] M. M. Mandurah, K. C. Saraswat and T. I. Kamins, "A Model for Conduction in Polycrystalline Silicon, Part II: Comparison of Theory and Experiment," _IEEE Trans. Electron Dev._, ED-33, 1171 (1981).

[4.29] E. A. Irene and D. W. Dong, _J. Electrochem. Soc._, 129, 1347 (1982).
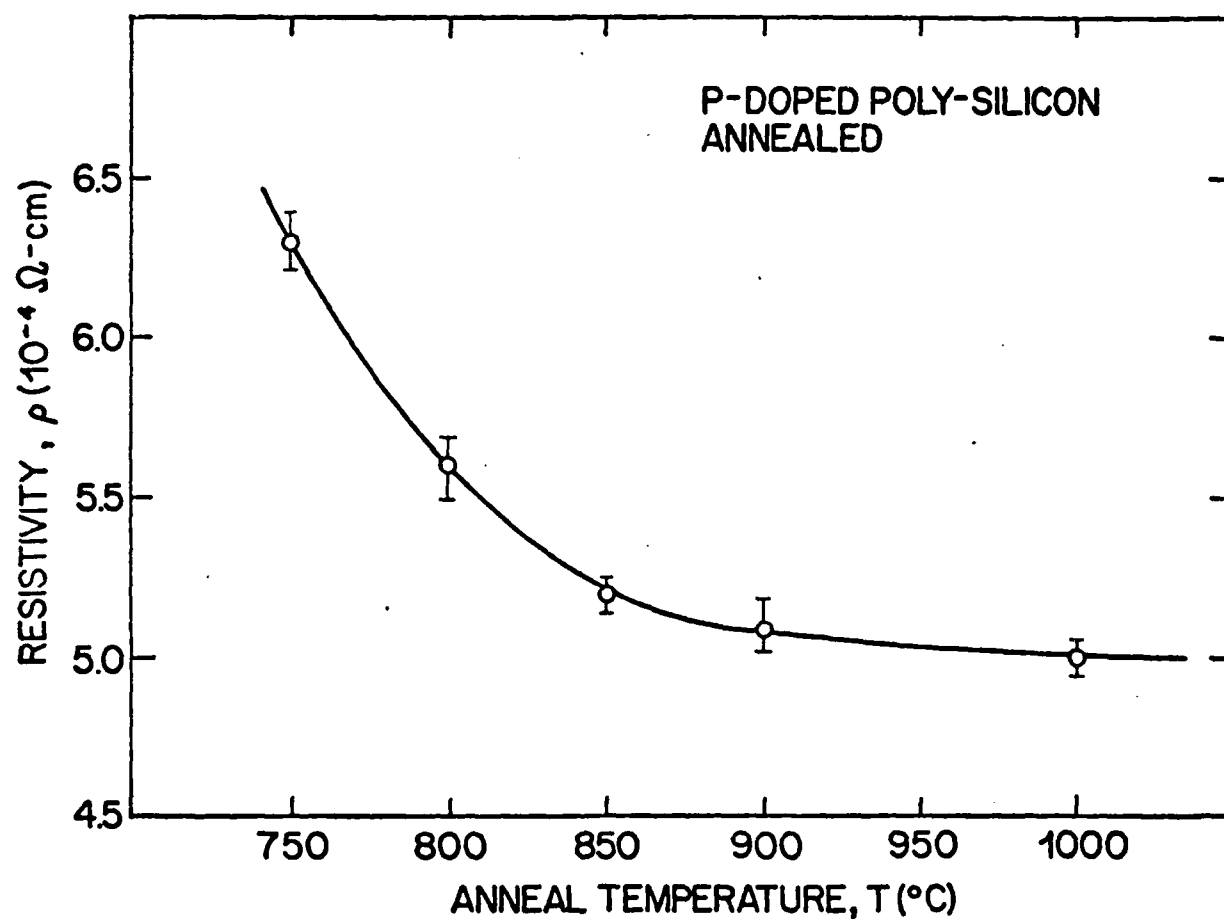
Fig. 4.1:   Resistivity of phosphorus doped poly-silicon films after annealing
            in a nitrigen ambient at 750°, 800°, 850°, 900°, and 1000°C for 140,
            64, 24, 13, and 2 hr., respectively.
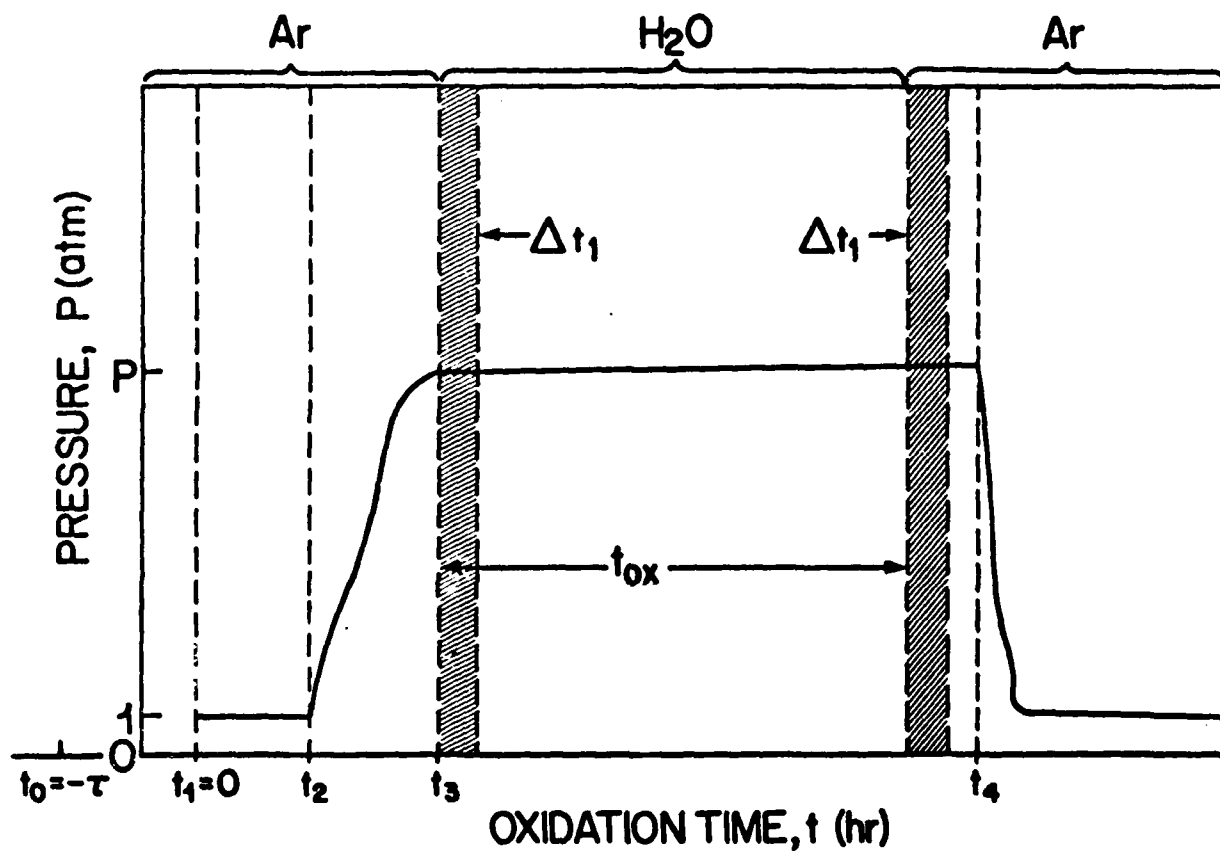
Fig. 4.2: Typical high pressure oxidation cycle used in the oxidation of phosphorus doped poly-silicon films in pyrogenic H_2O. The various stages of oxidation and the change in ambients are shown.
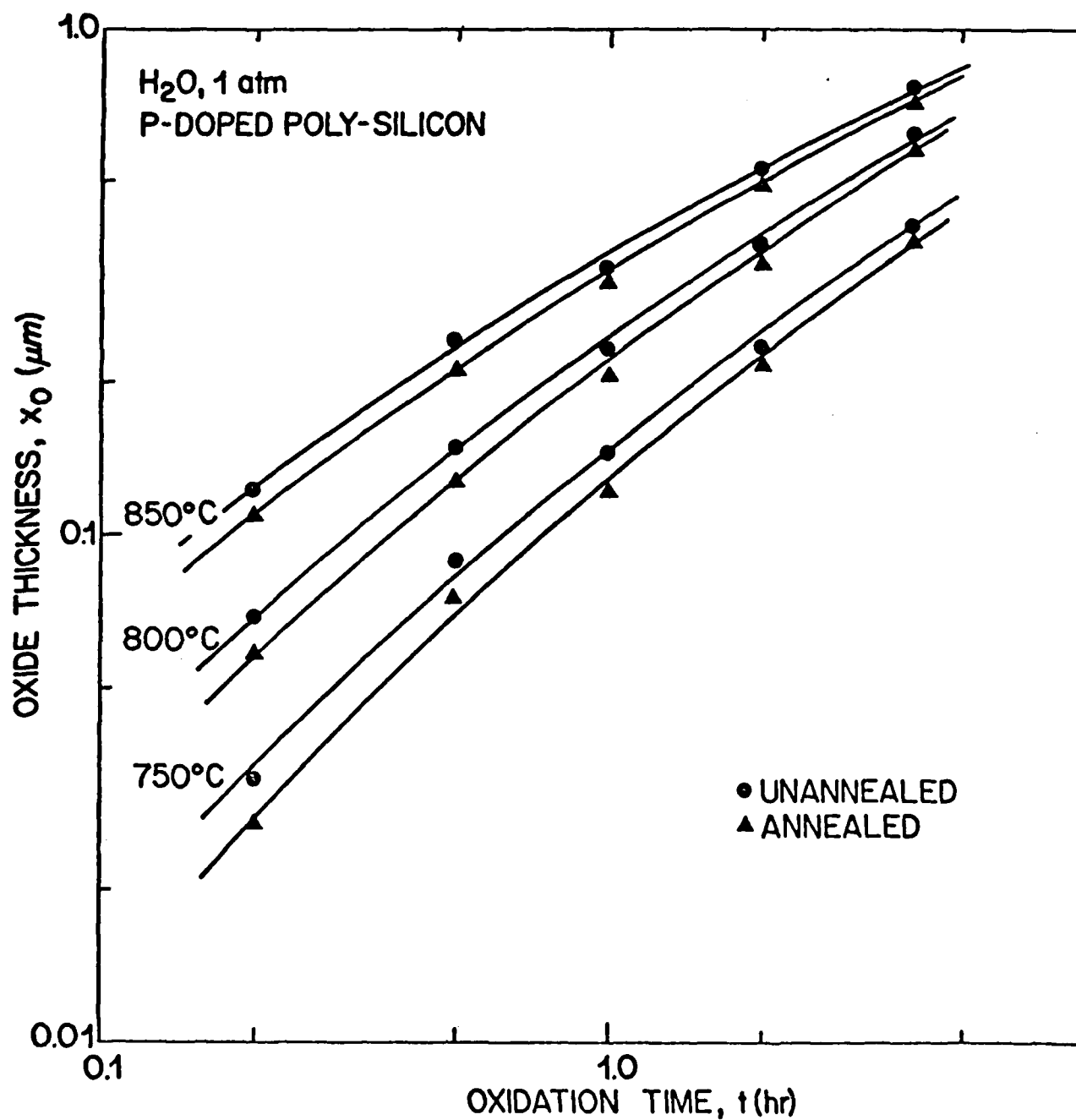
Fig. 4.3: Oxide thickness versus oxidation time for annealed and unannealed phosphorus doped poly-silicon films oxidized in pyrogenic H_2O at 1 atm and 750° – 850°C.
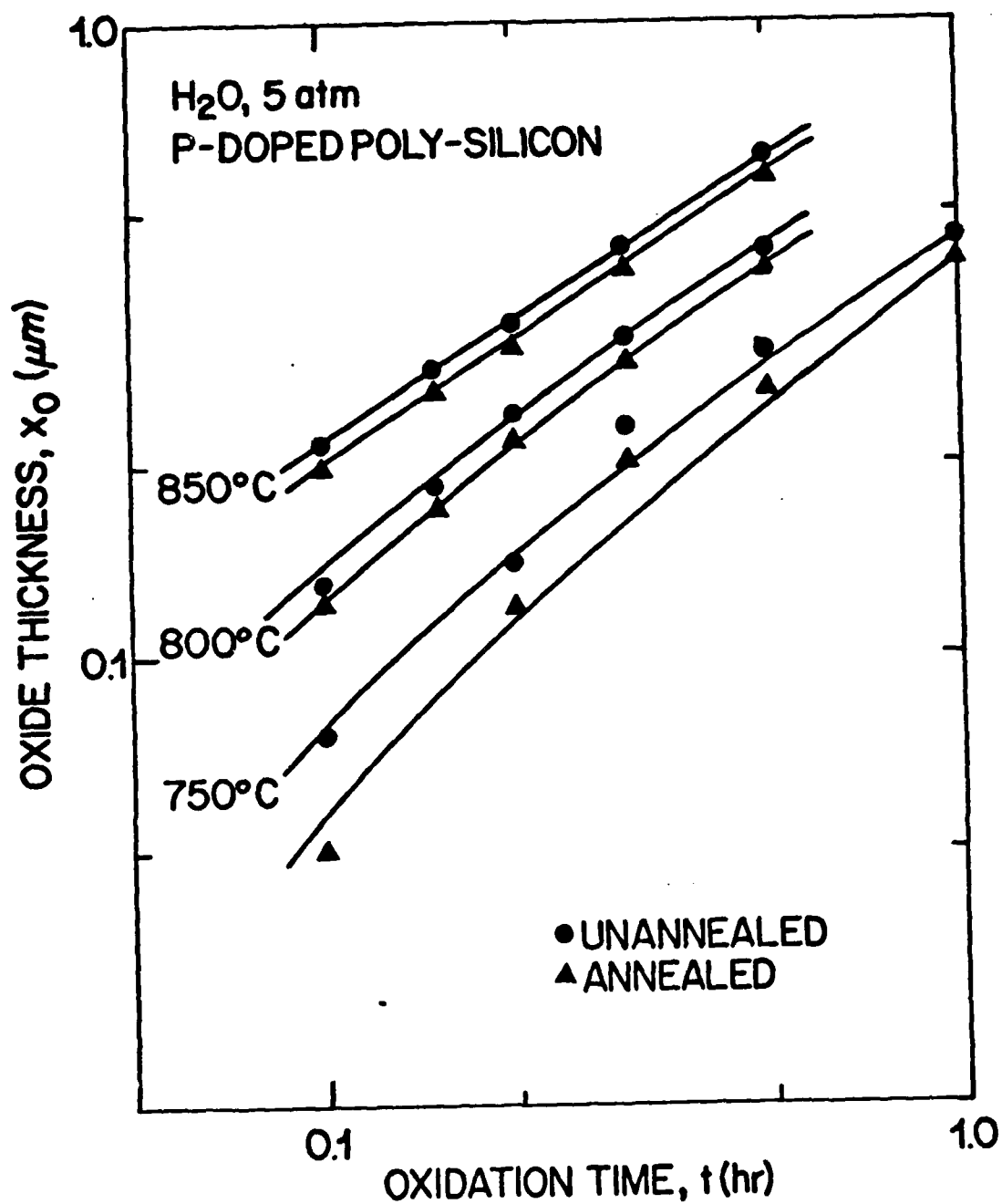
Fig. 4.4: Oxide thickness versus oxidation time for annealed and unannealed phosphorus doped poly-silicon films oxidized in pyrogenic H_2O at 5 atm and 750° - 850°C.

174
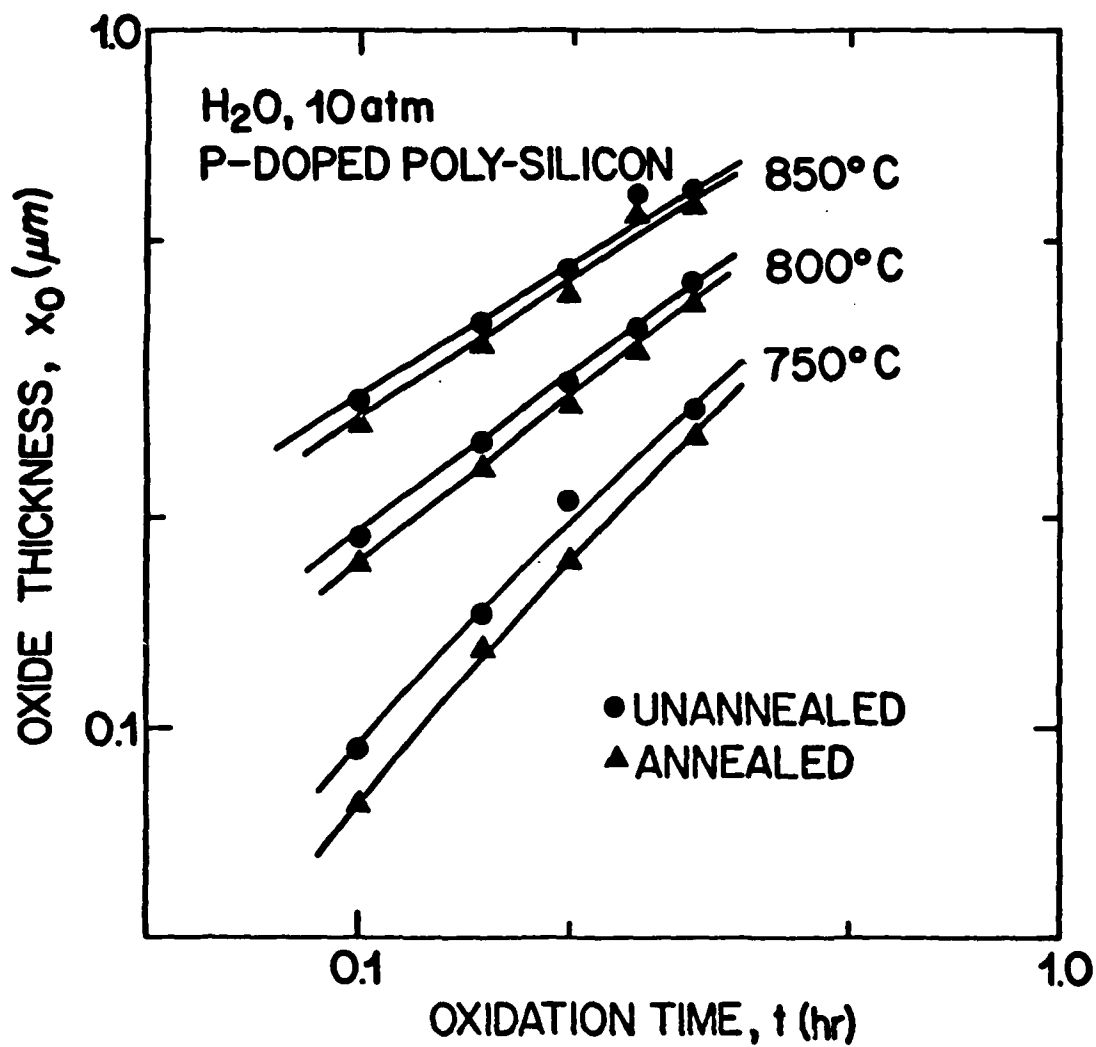
Fig. 4.5: Oxide thickness versus oxidation time for annealed and unannealed phosphorus doped poly-silicon films oxidized in pyrogenic H_2O at 10 atm and 750° - 850°C.
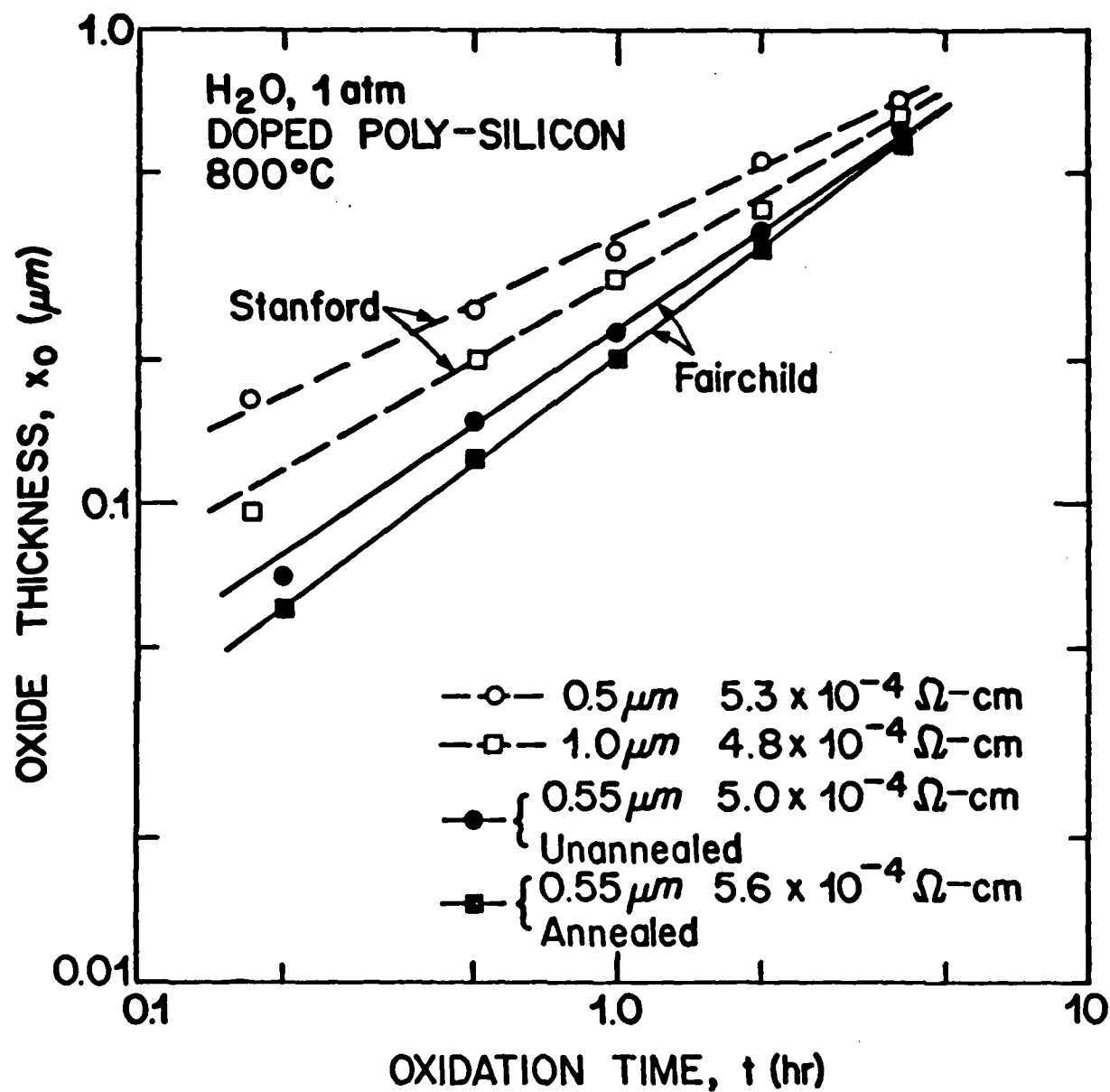
Fig. 4.6:   Oxide thickness versus oxidation time for poly-silicon films prepared at Fairchild and at Stanford. The phosphorus doped poly-silicon films were oxidized in pyrogenic H_2O at 1 atm and 800°C. The initial polysilicon film thicknesses and resistivities are indicated.
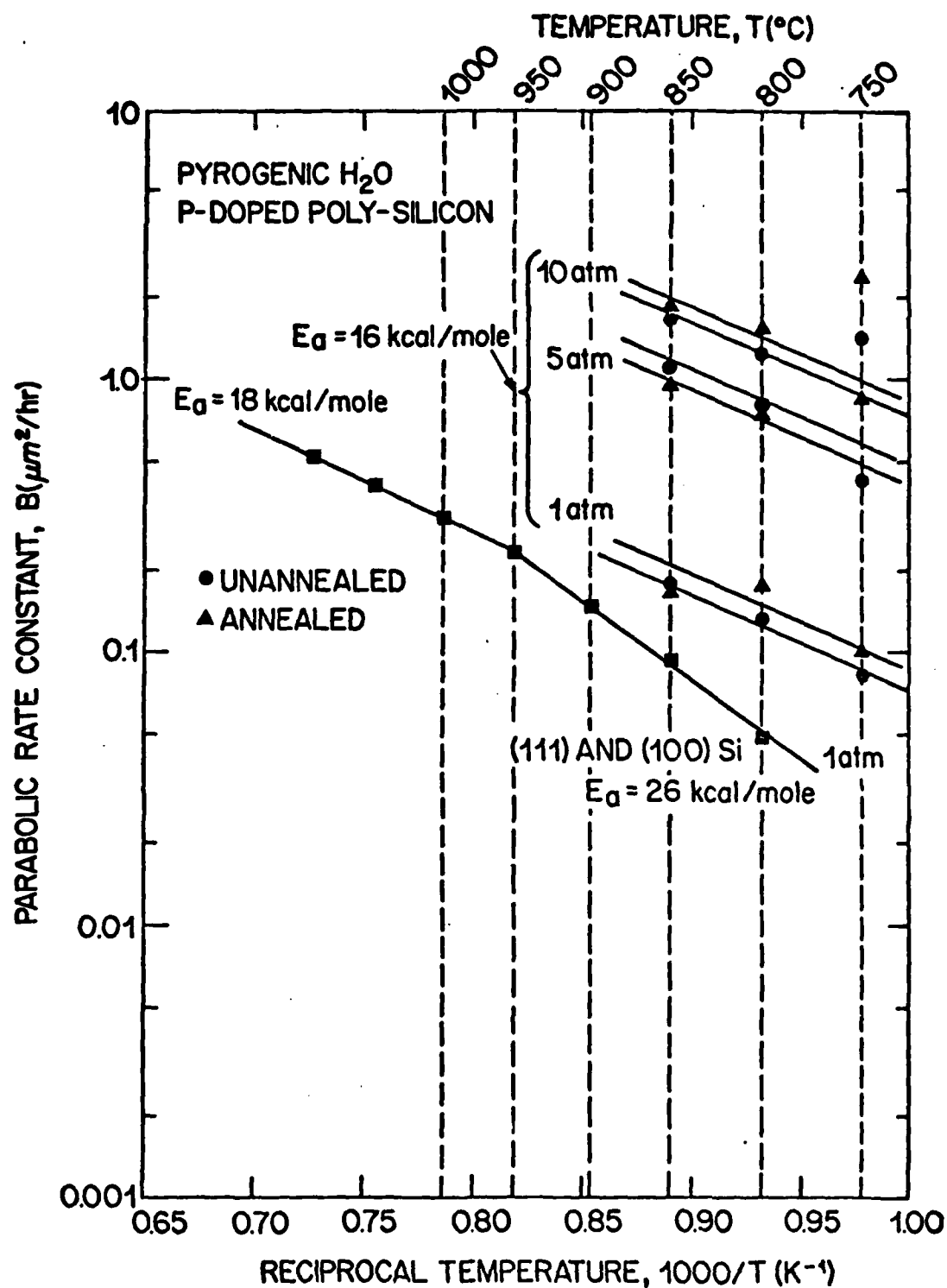
Fig. 4.7: Parabolic rate constant B versus 1000/T for annealed and unannealed phosphorus doped poly-silicon films oxidized in pyrogenic H_2O at 1, 5, and 10 atm over the temperature range of 750° - 850°C.
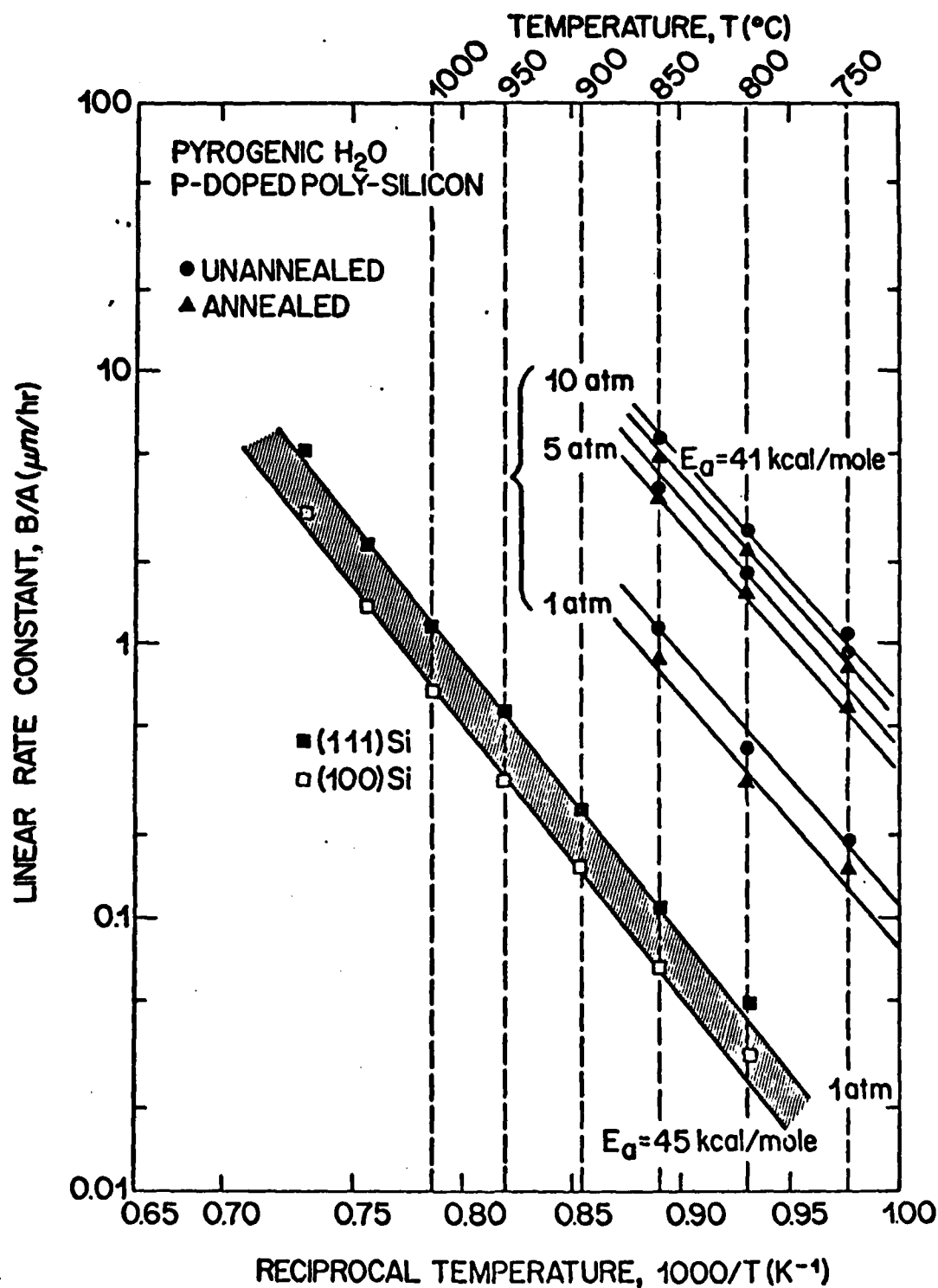
Fig. 4.8:   Linear rate constant B/A versus 1000/T for annealed and unannealed
phosphorus doped poly-silicon films oxidized in pyrogenic H_2O at
1, 5, and 10 atm over the temperature range of 750° - 850° C.
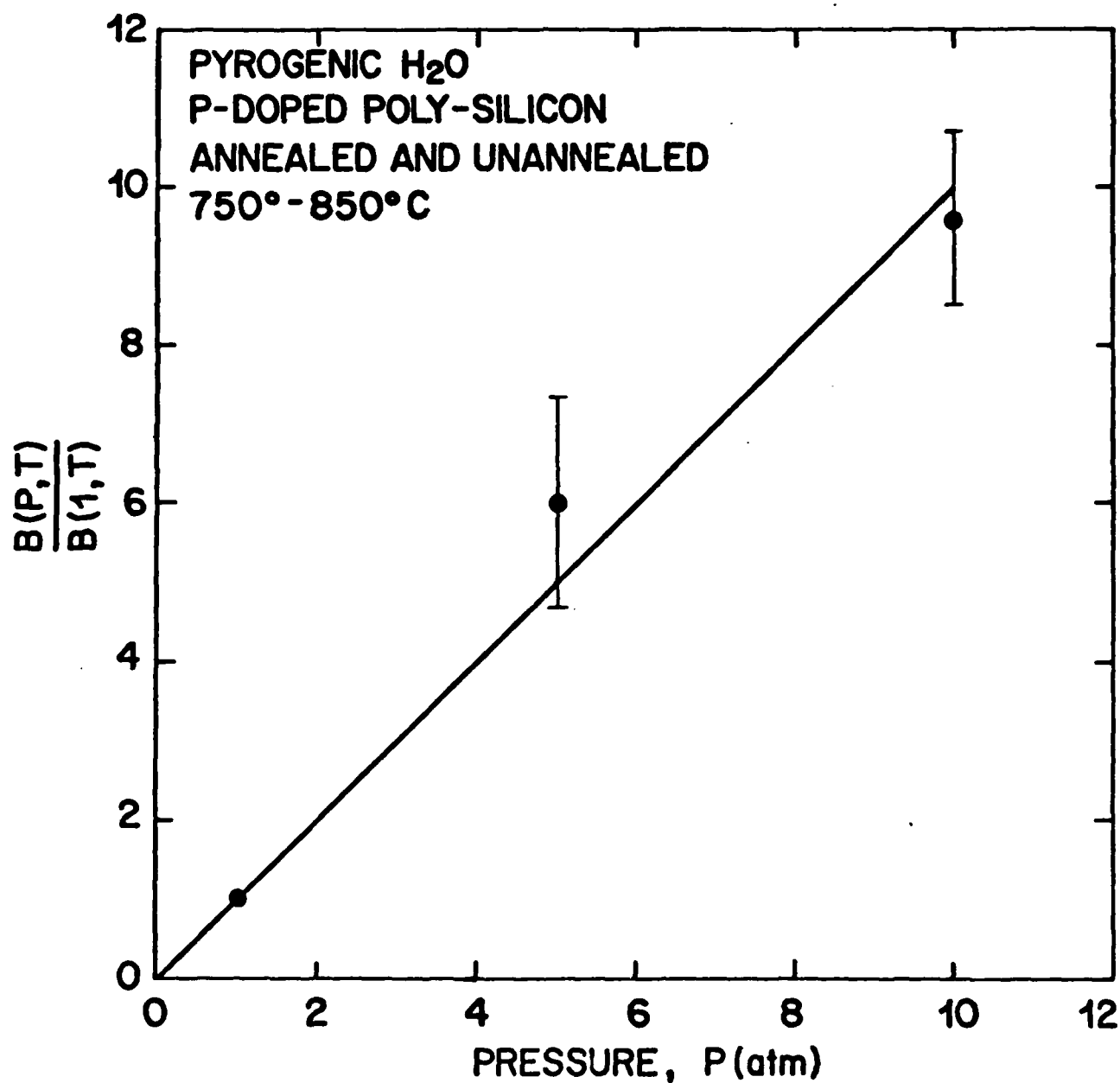
Fig. 4.9:   Normalized parabolic rate constant $B(P,T)/B(1,T)$ versus H_2O pressure for annealed and unannealed phosphorus doped poly-silicon films oxidized in pyrogenic H_2O at 1, 5, and 10 atm. Temperature range was 750° - 850°C.
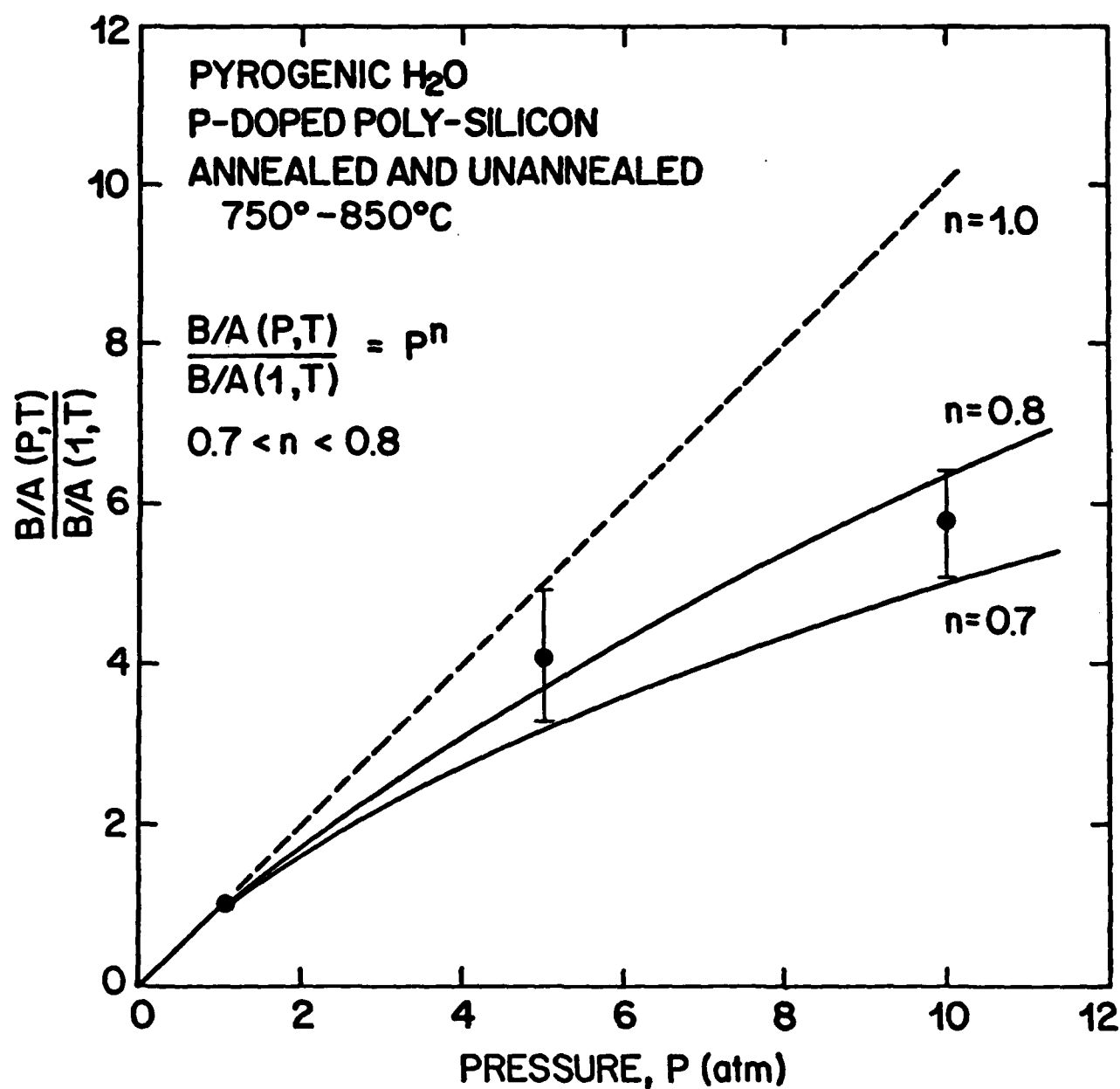
Fig. 4.10: Normalized linear rate constant [B/A(P/T)]/[B/A(1,T)] versus H_2O pressure for annealed and unannealed phosphorus doped polysilicon films oxidized in pyrogenic H_2O at 1, 5, and 10 atm. Temperature range was 750° - 850°C.
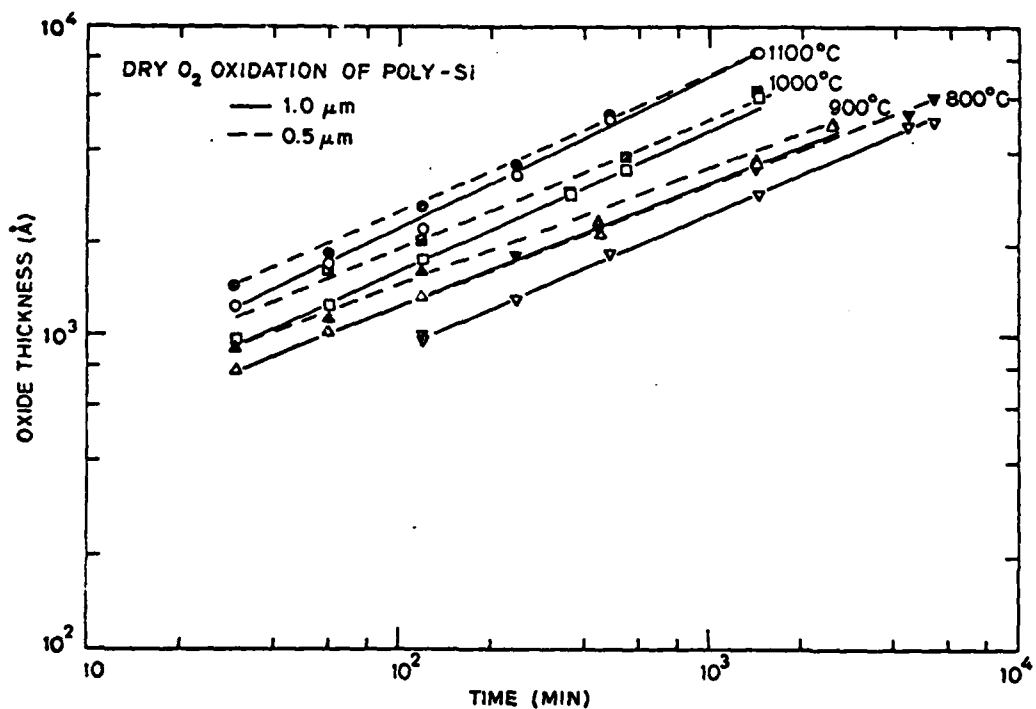
Fig. 4.11a: Oxide thickness vs. oxidation time for dry O_2 oxidation of phposphorus doped polysilicon films.
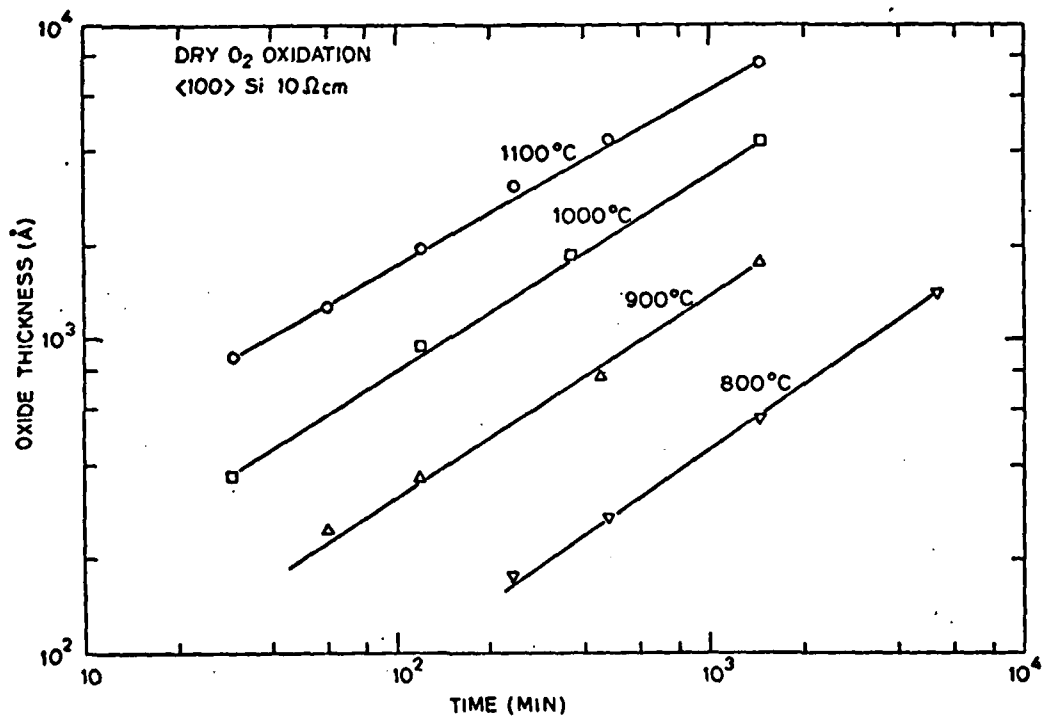
Fig. 4.11b: Oxide thickness vs. oxidation time for dry O_2 oxidation of lightly doped (100) silicon.
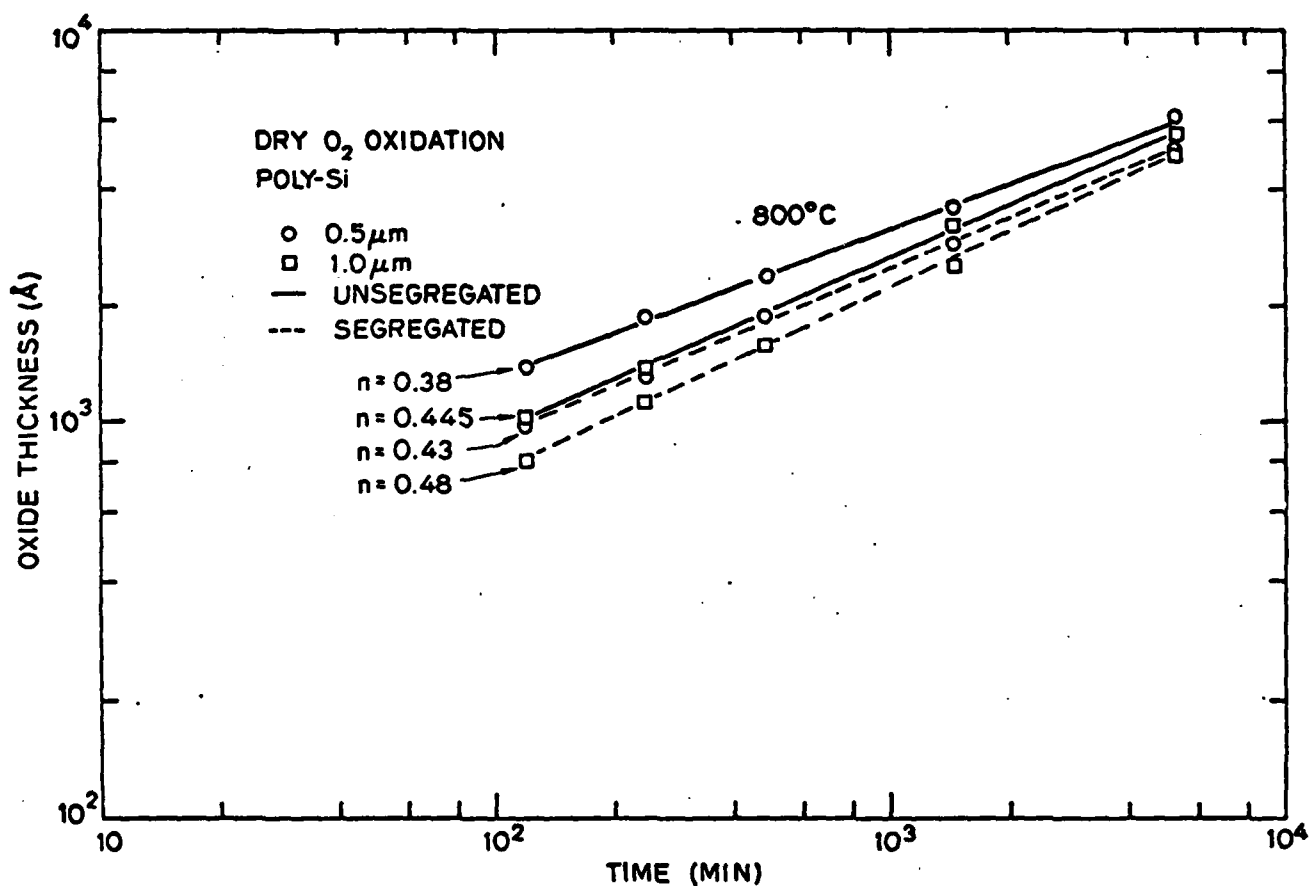
Fig. 4.12: Oxide thickness vs. oxidaton time for wet O_2 (O_2 bubbled through 95°C H_2O) oxidation of phosphorus-doped polysilicon films and lightly doped (100) silicon.

182
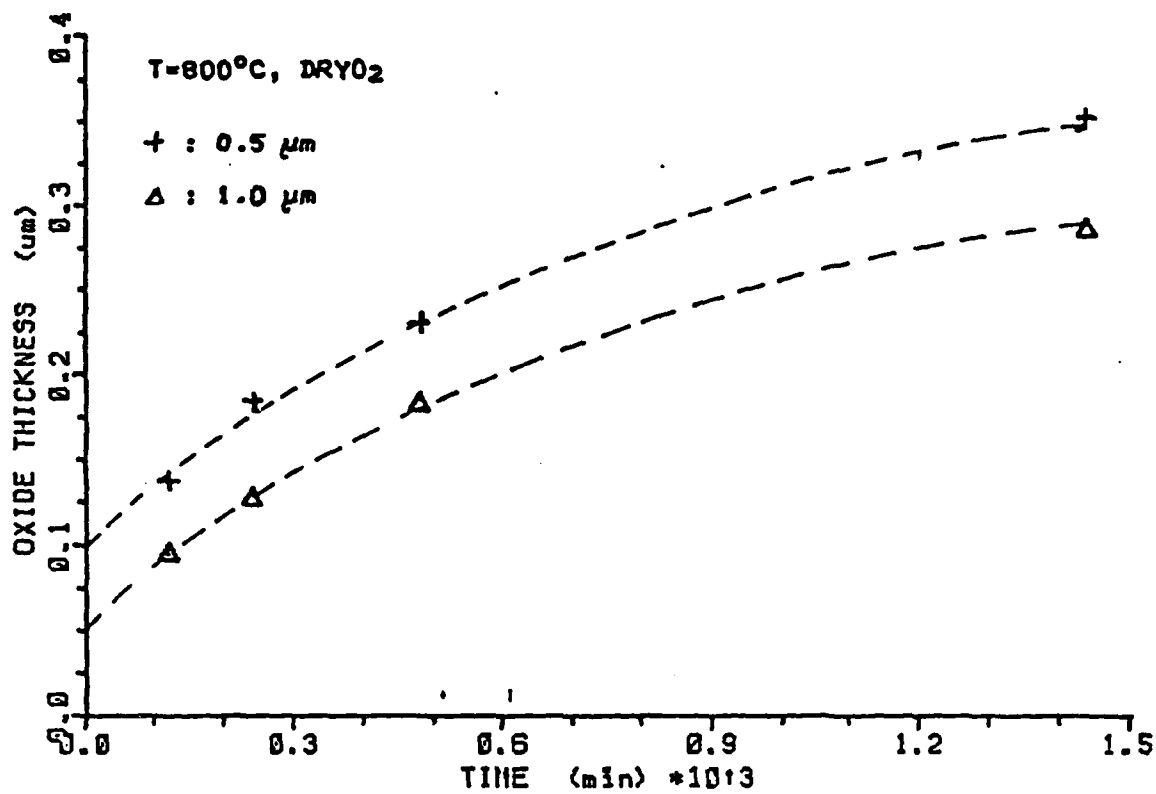
Fig. 4.13: A linear plot of oxide thickness vs. time to extract X_i.
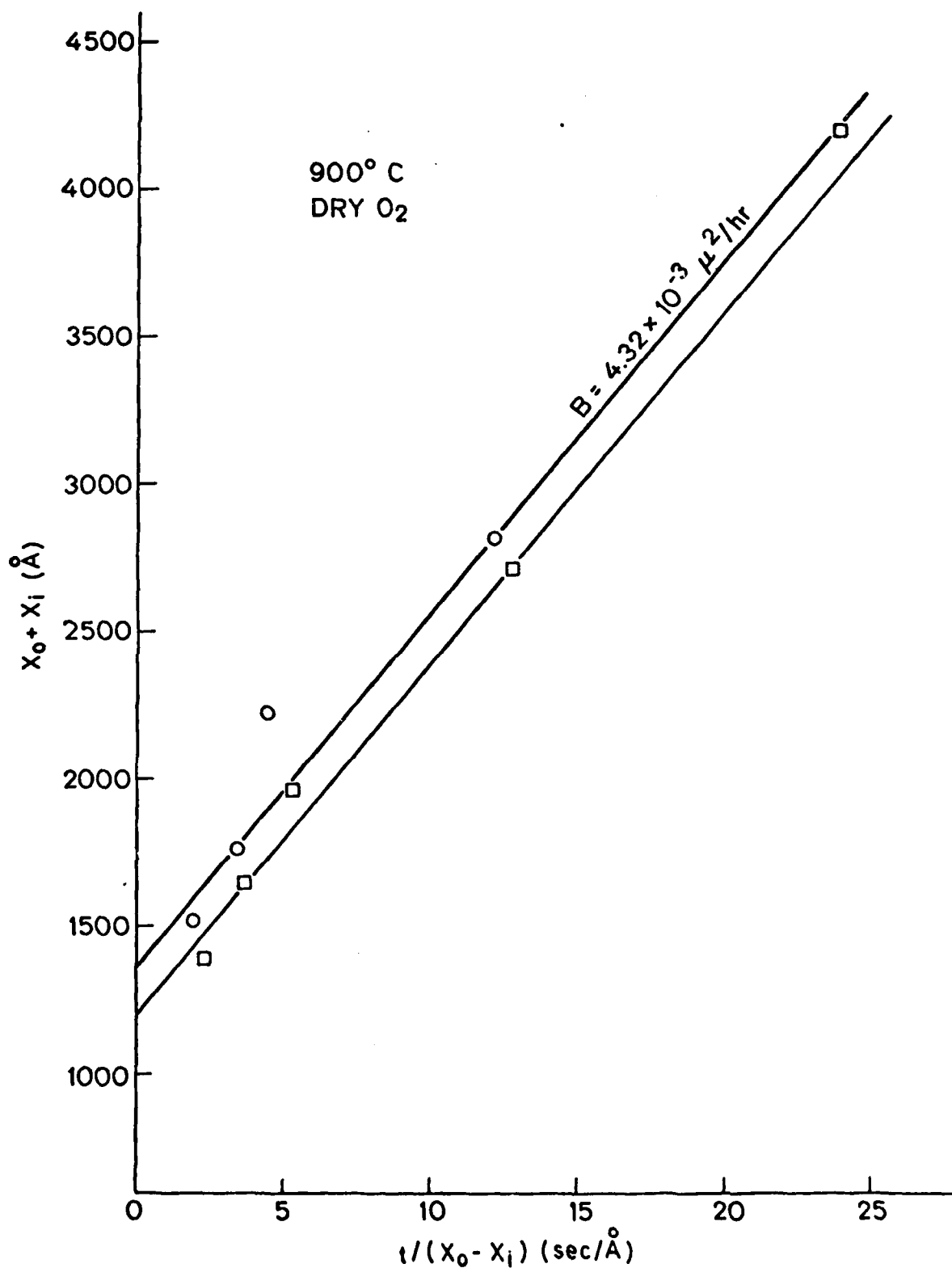
Fig. 4.14: $(X_0 + X_i)$ vs. $t/(X_0 - X_i)$ plot to extract B and A.
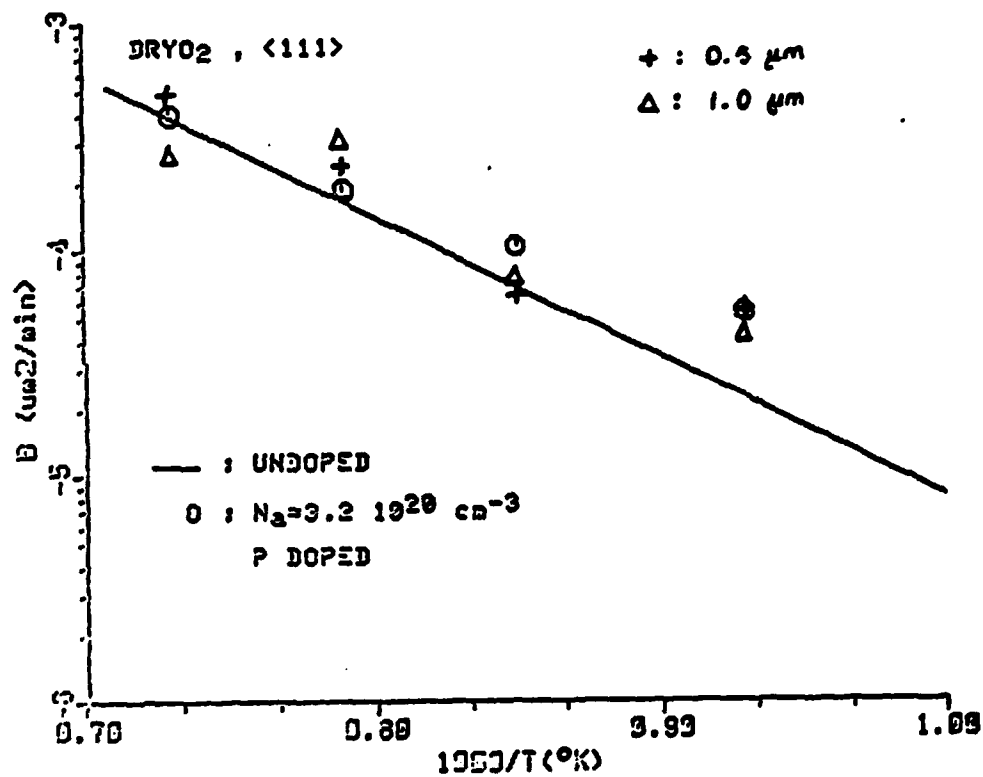
Fig. 4.15: Parabolic rate constant as a function of inverse of oxidation temperature.

Fig. 4.16a

Fig. 4.16b

186
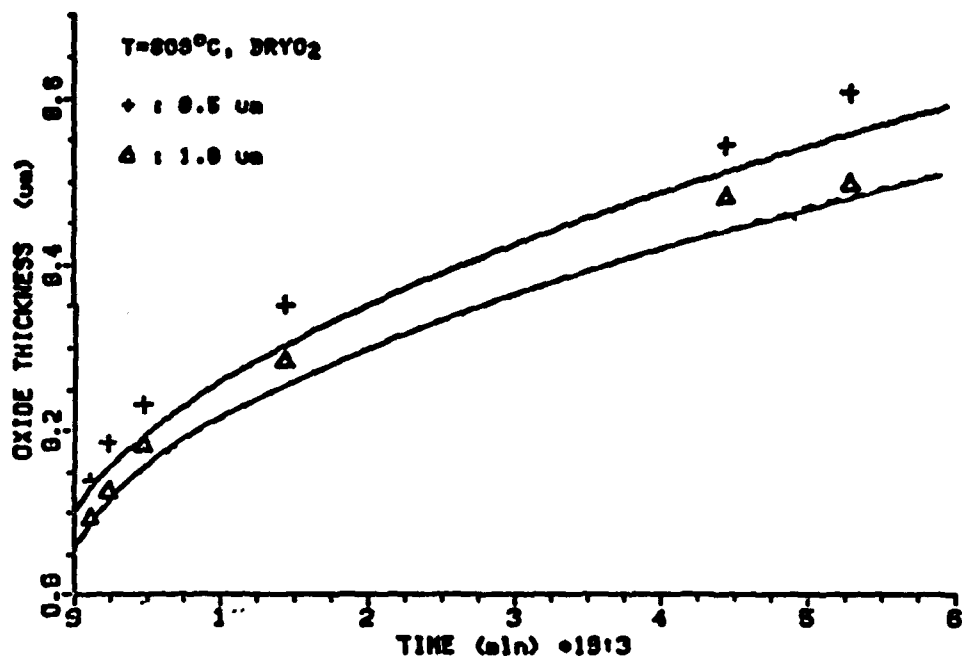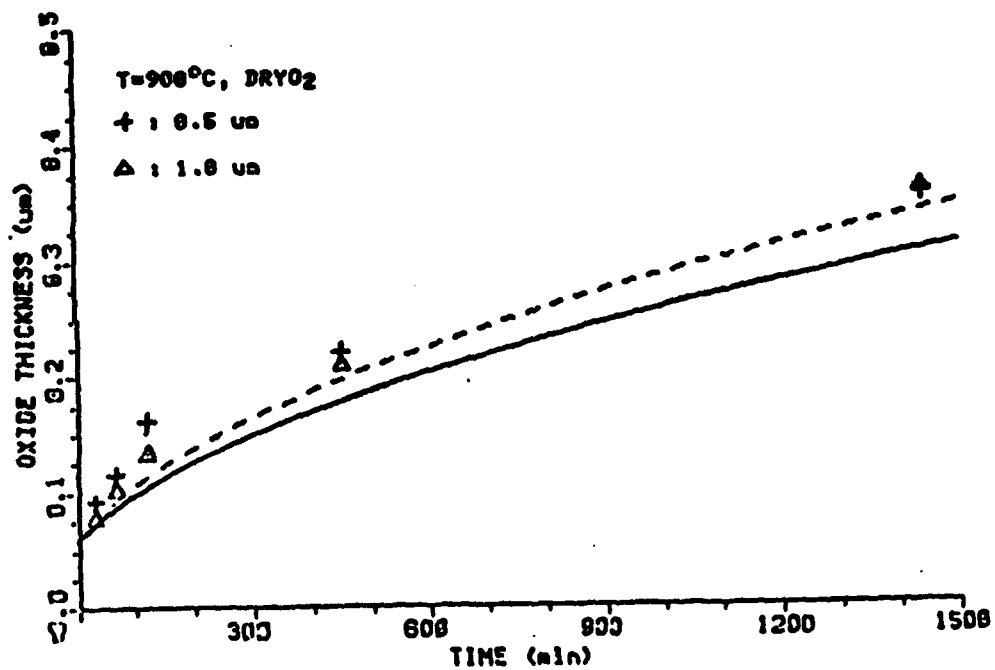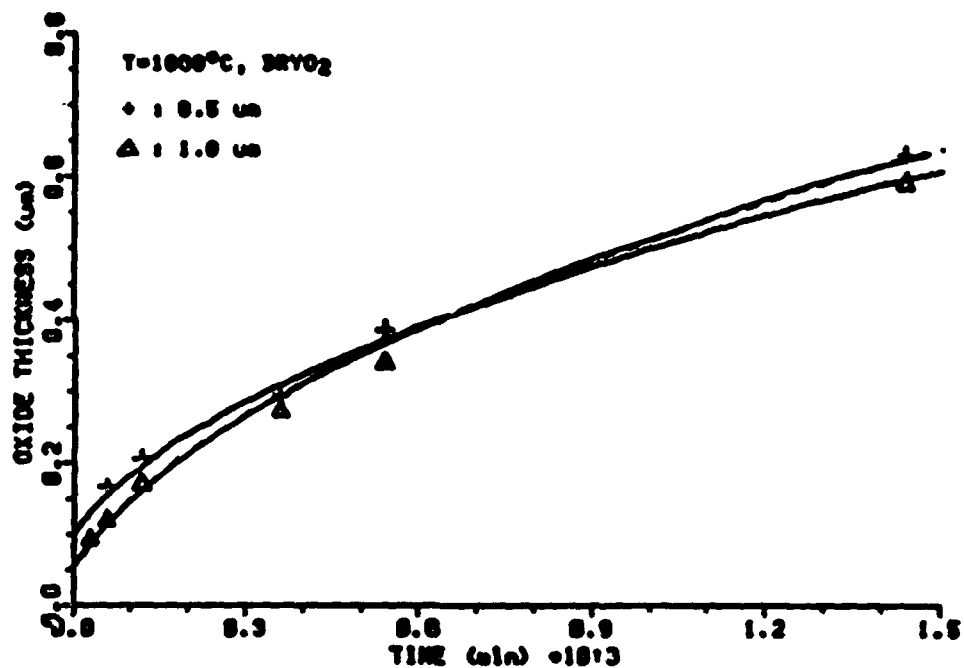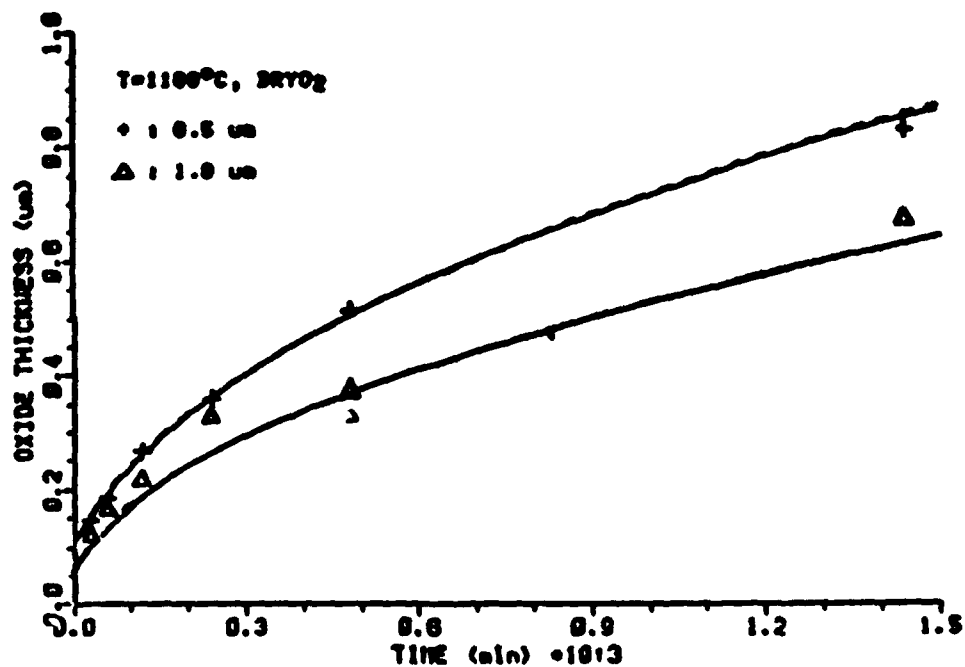
Fig. 4.16c

Fig. 4.16d

Fig. 4.16: Comparison of experiment (data points) and theoretical calculations (solid lines) using the Deal-Grove model (a) 800°C, (b) 900°C, (c) 1000°C, (d) 1100°C.

187
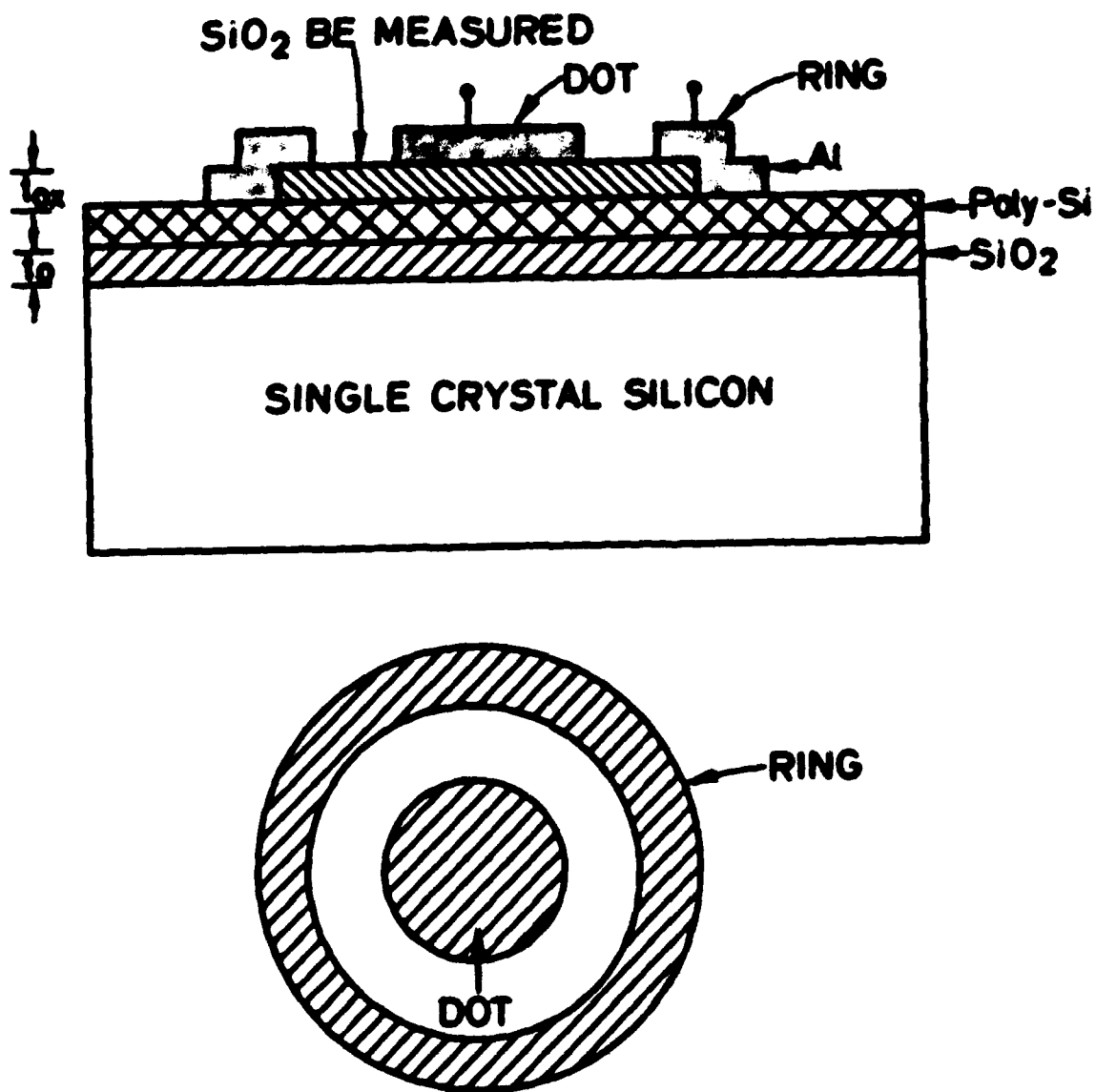
Fig. 4.17: The ring-dot structure for C-V measurements.
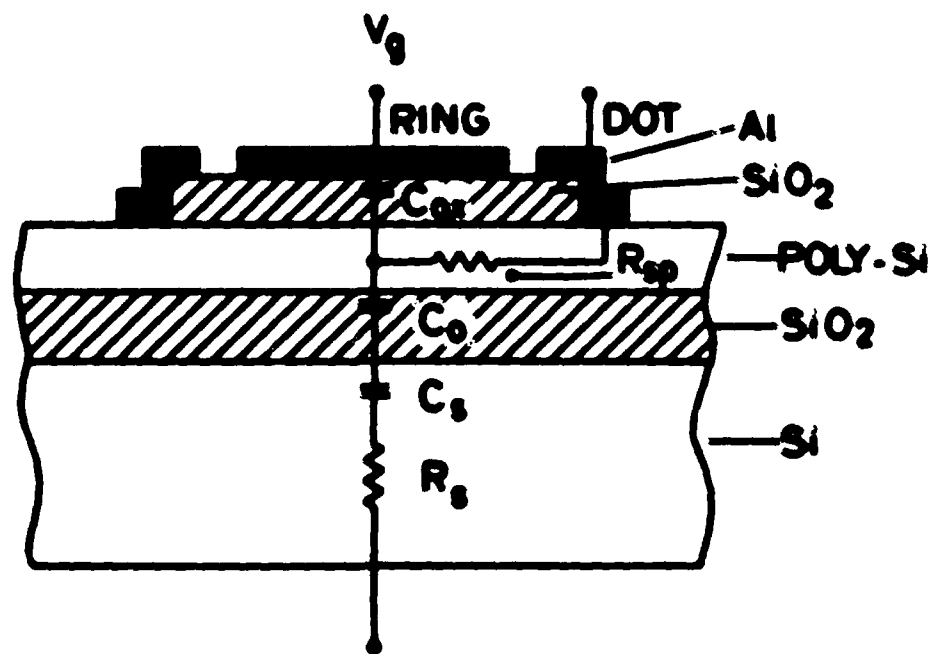
Fig. 4.18: Parasitics associated with the ring-dot structure.
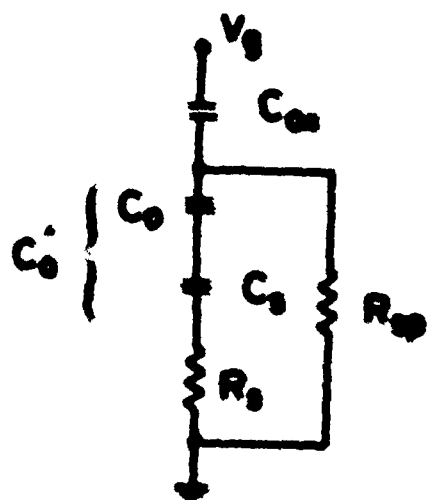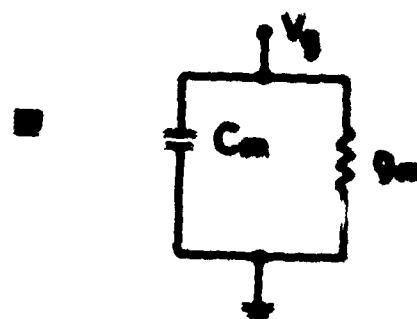
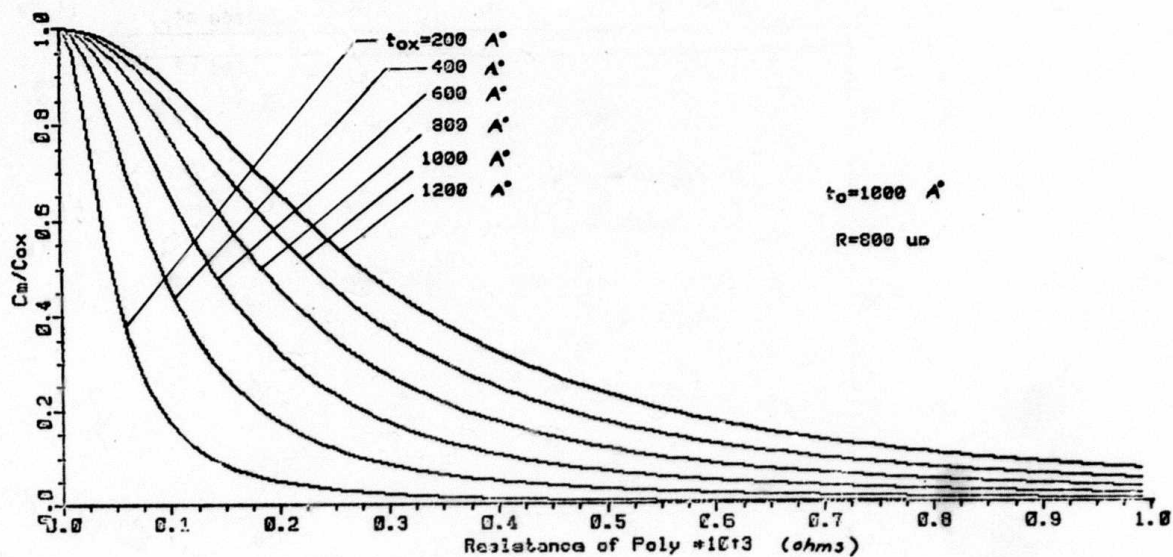Fig. 4.19: (a) lumped element equivalent circuit, and (b) measured equivalent of the ring-dot structure.
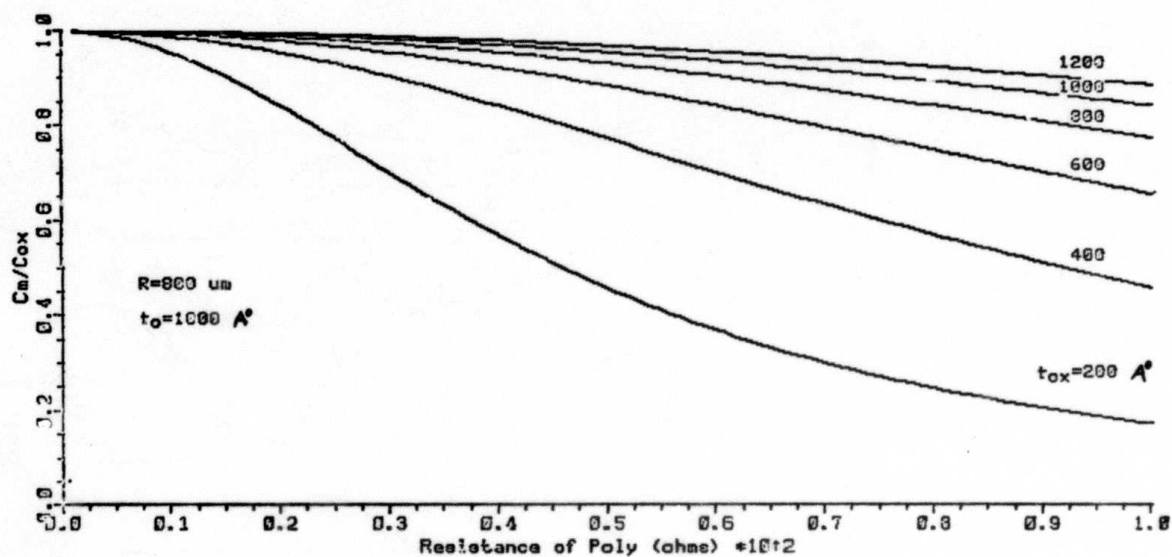
Fig. 4.20: C_m/C_{ox} as a function of poly-Si resistance R_{sp}, (a) R_{sp} up to 10^3 ohms, (b) R_{sp} up to 10^2 ohms.
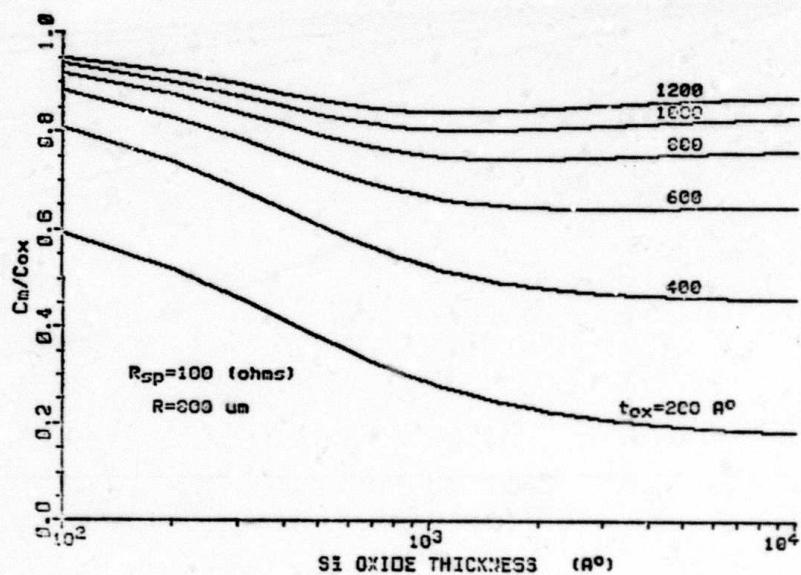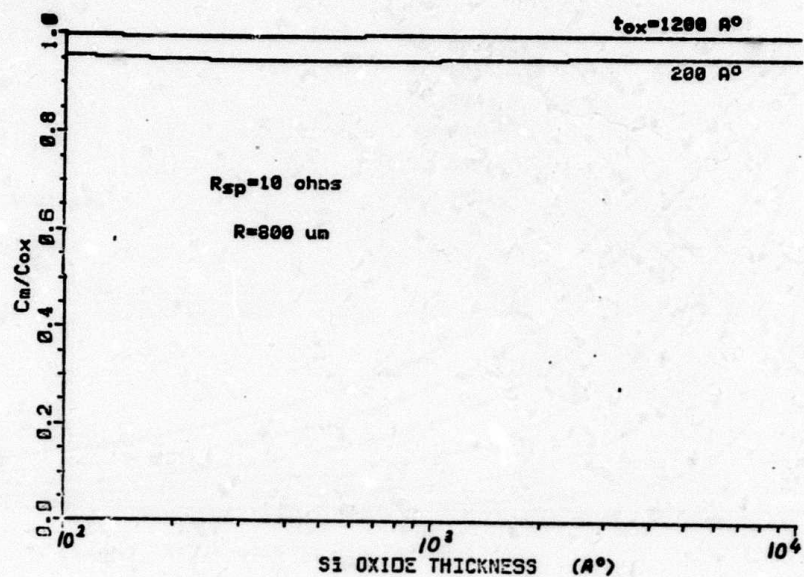
191

Fig. 4.21: C_m/C_{ox} vs. oxide thickness on single crystal silicon for (a) R_{sp} = 10 ohms, (b) R_{sp} = 100 ohms.
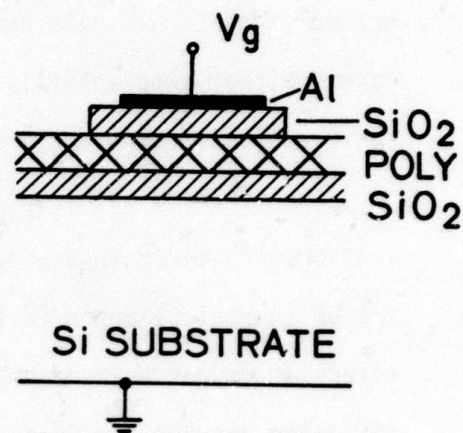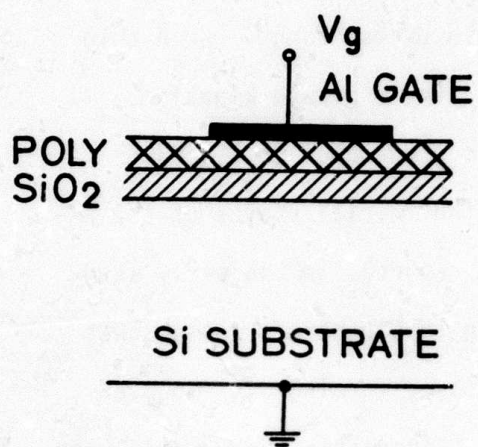
192

Fig. 4.22: (a) A novel structure to measure the parasitics of (b).

193

5.   THIN OXIDE GROWTH KINETICS IN DRY OXYGEN

H. Z. Massoud, C. P. Ho, J. D. Plummer

5.1   INTRODUCTION

Very thin ($\leq$ 500 Å) layers of thermally grown SiO_2 are becoming essential components of VLSI circuits. State of the art high-performance MOS integrated circuits use gate oxides in the 250-400 Å thickness range. Such thin layers differ substantially from thicker layers in their growth kinetics, their optical, structural, dielectric and electric properties [5.1,5.2].

The enhanced oxidation process occurring in the initial stages of the oxidation of silicon in dry oxygen was observed and reported in the early studies of thermal oxidaton [5.1]. Deal and Grove, in that paper, proposed that molecular oxygen is ionized at the outer SiO_2 surface, forming O_2^-. Coupled diffusion between the O_2^- and holes then effectively increases the O_2^- diffusion rate until an oxide thickness on the order of the extrinsic Debye length is reached. Revesz and Evans [5.3] and Ing et. al [5.4] proposed that SiO_2 may contain structural micro-heterogeneities, especially channels (<50 Å in diameter), along which transport of the diffusing species might occur preferentially. Irene [5.5] suggested that micropores in the SiO_2 film provide a short circuit path to the $Si-SiO_2$ interface for oxidant species which do not attack SiO_2 (such as O_2 related oxidants). Derbenwick and Anderson [5.6] proposed that an increased solubility of O_2 in SiO_2 for very thin layers might be important in describing the rapid initial oxidation in dry oxygen.

The growth kinetics of silicon in dry oxygen were successfully modelled by Deal and Grove [5.1] using a simple linear-parabolic kinetics relationship. In this model oxygen molecules diffuse through the SiO_2 layer already existing and react at the $Si-SiO_2$ interface (to form SiO_2) by breaking Si-Si bonds. The oxidation rate is expressed as:

$$\frac{dX_{ox}}{dt} = \frac{B}{2X_{ox} + A} \tag{5.1}$$

and the relationship between the oxide thickness X_{ox} and the oxidation time t_{ox} can be written as:

$$(X_{ox}^2 - X_i^2) + A(X_{ox} - X_i) = Bt_{ox} \tag{5.2a}$$

or

$$X_{ox}^2 + AX_{ox} = B(t_{ox} + \tau) \tag{5.2b}$$

where B and B/A are the parabolic and linear rate constants respectively, X_i is the value of the oxide thickness at $t = 0$, and τ is a correction parameter to account for the presence of any initial oxide X_i ($\tau = (X_i^2 + AX_i)/B$).

This model relates the rate constants B and B/A to the physico-chemical constants of the process described above and models the oxidation data beyond 350 Å fairly accurately. However, it predicts a linear dependence of the rate constants on the partial pressure of oxygen in the oxidizing ambient. This is in contradiction with the findings of van der Meulen and Ghez [5.7,5.8]. Their work involved oxidations between 700°C and 1000°C, in dry O_2 and in O_2/N_2 mixtures with the O_2 partial pressure as low as 0.01 atm. They observed that the linear rate constant depends on oxygen partial pressure with a $p^{0.5}$ dependence at 700°C varying to a $p^{1.0}$ dependence at 1000°C.

In an effort to explain this behavior, molecular oxygen was assumed to decompose into atomic oxygen at the $Si-SiO_2$ interface. The surface reaction then proceeds via both molecular and atomic oxygen, with the molecular reaction dominating at high temperature (resulting in the $p^{1.0}$ pressure dependence) and with the atomic oxygen reaction dominating at low temperature (resulting in the $p^{0.5}$ pressure dependence). The decrease in activation energy for the interface

195

reaction at lower temperatures is consistent with our recent experimental data [5.9].

Recently Blanc [5.10] proposed a kinetic model for thin oxides which, in essence, assumes that the reaction involving molecular oxygen does not occur at all. Based upon this he derives a growth law given by

$$\frac{1}{2} u + \frac{1}{4} (\exp 2u - 1) = bt \qquad (5.3)$$

where $ax = \sinh u$ and a and b are rate constants. For thin oxides, this predicts a linear parabolic growth law but with different effective rate constants. Blanc compared his theoretical predictions with the data of Irene [5.5] at temperatures of 780° - 980°C with apparently good agreement. As Blanc points out, however, his model predicts a $p^{0.5}$ pressure dependence for the interface reaction rate; this would seem to be in disagreement with the data of van der Meulen and Ghez.

5.2 EXPERIMENTAL RESULTS AND PARAMETER EXTRACTION

Accurate modeling of the thin oxide regime has been most hindered by the lack of a comprehensive body of experimental data. Such data would help differentiate between the various proposed models or in defining a new model. A major effort in this program has, therefore, been aimed at obtaining such a body of data. Since an oxidation rate law may result from a number of possible mechanisms, the observed reaction rate law may not be conclusive evidence for a particular mechanism. Therefore, a large matrix of experimental conditions has been sought in order to reduce the number of possible mechanisms involved by discarding those which are valid only under some of the experimental conditions. Thus, the effect of temperature, partial pressure of oxygen in the oxidizing ambient, silicon orientation and the doping density of the silicon substrate were investigated. Oxidations were carried out in collaboration with E. A. Irene of IBM, using an in-situ ellipsometry capability and

196

covered the temperature range of 800-1000°C, partial pressures of 0.01-1.0 atmospheres of O_2 in argon, three substrate orientations and a variety of substrate doping levels.

In this report, we first review the methods used in extracting rate constants from the oxidation data. Accurate values of the rate constants are essential to accurately model the oxidation enhancement in the initial stages.

5.2.1. Extraction of Linear and Parabolic Rate Constants

In their description of silicon oxidation, using the linear-parabolic relationship, Deal & Grove [5.1] found that, when using oxygen bubbled through water at 95°C, the kinetics followed the relationship from the start of the oxidation (i.e. with only the native oxide present). However, when using dry oxygen (water content less than 5 ppm), the oxide growth kinetics did not extrapolate to zero initial thickness. This is illustrated in Fig. 5.1 (a and b).

A value of X_i = 230 ±30 Å was found in the temperature range 700 - 1200°C for the case of <111> silicon oxidized in dry oxygen. This value of X_i or the corresponding values of τ, were used to account for the fast initial oxidation process in this case. It should be noted (Fig. 5.1a) that the linear parabolic relationship only fits the data for oxides thicker than X_c or times longer than t_c.

5.2.2 Methods for Rate Constants Extraction

a. Plot $(X_{ox} + X_i)$ vs. $t_{ox}/(X_{ox} - X_i)$

By dividing both sides by $(X_{ox} - X_i)$, equation 5.2a can be rewritten as:

$$(X_{ox} + X_i) + A = B \frac{t_{ox}}{(X_{ox} - X_i)} \tag{5.4}$$

It is therefore seen that a plot of $(X_{ox} + X_i)$ vs. $t_{ox}/(X_{ox} - X_i)$ should result in a straight line. The slope of this line is B and the intercept at $t_{ox}/(X_{ox} - X_i) = 0$ is $-A$. This is illustrated in Fig. 5.2a.

b. Plot X_{ox} vs. $(t + \tau)/X_{ox}$

In this method, an initial value of X_i is adopted. The corresponding value of τ is estimated graphically by plotting X_{ox} vs. t_{ox} and extrapolating the curve back through X_i to the time axis (Fig. 5.1a).

By dividing both sides by X_{ox}, equation 5.2b may be rewritten as:

$$X_{ox} + A = \frac{B(t + \tau)}{X_{ox}} \tag{5.5}$$

A plot of X_{ox} vs $(t + \tau)/X_{ox}$ should therefore yield a straight line where the slope is again B and the intercept at $(t + \tau)/X_{ox} = 0$ is $-A$. This is shown in Fig. 5.2b. It is easily seen that the accuracy with which the rate constants are calculated depends entirely on how good the estimate of X_i or τ is. One solution to this difficulty is, instead of estimating a thickness X_i at $t_{ox} = 0$, we could pick a thickness X_1 at $t_{ox} = t_1$ where X_1 is an oxide thickness beyond the fast initial oxidation regime. This is equivalent to shifting the time axis by t_1. This is illustrated in Fig. 5.3.

For any thickness X_j and time t_j we can write:

$$(X_j^2 - X_1^2) + A(X_j - X_1) = B(t_j - t_1) \tag{5.6}$$

A plot of $(X_j + X_1)$ vs. $\dfrac{t_j - t_1}{X_j - X_1}$ could then be used to calculate B and B/A. X_i or τ could then be calculated by substituting in eq. 5.2a or 5.2b.

c. Plot $1/(dX_{ox}/dt)$ vs. X_{ox}

In this method, the oxidation rate as a function of the oxide thickness X_{ox} is obtained either graphically or numerically. The inverse of the rate

is then plotted as a function of thickness. Equation 5.1 is rewritten as

$$\frac{1}{\frac{dX_{ox}}{dt}} = \left(\frac{2}{B}\right) X_{ox} + \frac{A}{B} \qquad (5.7)$$

Therefore a plot of the inverse of the oxidation rate vs. thickness should yield a straight line. The slope of this line is (2/B) and the intercept at $X_{ox} = 0$ is (A/B). Fig. 5.2C is a typical plot for this method.

In the inverse rate method, however, the accuracy of extracted rate constants is directly proportional to that of the differentiation process. It should be noted here that the closer the X_{ox} - t_{ox} data points are, the more accurate the calculation of the oxidation rate constants will be.

In the case of oxidation of <111> oriented silicon in dry oxygen, Deal and Grove [5.1] used the second method to calculate the rate constants. From the different values of τ for temperatures in the 700° - 1200°C range, corresponding values for X_i were obtained and a constant value of 230 A was adopted. The value of X_i used in the extraction of the rate constants has not been constant in the literature over the years. Table 1 shows the different values adopted over the span of 15 years.

TABLE 5.1   HISTORY OF X_i

Reference	Value of X_i
Deal & Grove, Ref. [5.1]	230 ± 30 A
Hess & Deal, Ref. [5.11]	160 ± 40 A
Razouk, Lie, & Deal, Ref. [5.9]	40 ± 0 A

5.2.3. Effect of X_i or τ on the Value of B & B/A:

The value of X_i adopted in this study was obtained by maximizing the number of (X_{ox}, t_{ox}) data points that fit a linear-parabolic relationship. Henceforth, this value will be referred to as the optimum X_i. This value of X_i was obtained by plotting $(X_{ox} + X_i)$ vs. $t_{ox}/(X_{ox} - X_i)$ for different values of X_i. Fig. 5.4 shows a typical example.

It was found that for values of X_i close to the optimum value, the data could be fit to a straight line down to a thickness X_c. X_c was found to be a sharp function of X_i. For data below X_c, the plot deviated sharply from a straight line due to the fast initial regime.

It is important to note that if the data correspond to oxide thicknesses greater than X_c, a straight line $(X_{ox} + X_i)$ vs. $t_{ox}/(X_{ox} - X_i)$ plot is obtained for a range of values of X_i, smaller and larger than X_i (opt). In Fig. 5.4 it is observed that:

- If $X_i = X_{i(a)} > X_{i(opt)}$, B is larger and (B/A) is smaller than the rate constants obtained with the opt. X_i (Fig. 5.4a). The linear-parabolic relationship fits the data beyond $X_{c(a)}$ where $X_{c(a)}$ > $X_{c(opt)}$ (Fig. 5.4b).

- If $X_i = X_{i(b)} < X_{i(opt)}$, B is smaller and (B/A) is larger. The linear-parabolic relationship and the $X_{ox} - t_{ox}$ data intersect at point P (Figs. 5.4a and 5.4b). The thickness beyond which the linear-parabolic relationship fits the data $X_{c(b)}$ was also found to be larger than that in the optimum case $X_{c(opt)}$ (Fig. 5.4b).

To estimate the error in the rate constants caused by improper values of X_i, a set of $X_{ox} - t_{ox}$ data were generated using the following expression to simulate the initial fast regime:

$$(X_{ox}^2(t) - X_{ox}^2(o)) + A(X_{ox}(t) - X_{ox}(o))$$

$$= Bt + K_1\tau_1 \left(1 - e^{-t/\tau_1}\right) + K_2\tau_2 \left(1 - e^{-t/\tau_2}\right) \qquad (5.8)$$

The values of B, B/A, $X_{ox}(o)$, K_1, τ_1, K_2 and τ_2 used were the values obtained experimentally from the $\langle 100 \rangle$ - 1000°C - 100% O_2 data in simulating the initial regime. These were:

B = 28590 A^2/min

B/A = 8.65 A/min

The oxidation rate was obtained by numerical differentiation.

An inverse rate plot resulted in a straight line (beyond the fast initial regime as shown in Fig. 5.5). It is seen that the linear-parabolic relationship applies beyond 400 A. The rate constants obtained from the straight line fit were identical to that used in the generation of the X_{ox} - t_{ox} values. Using the optimum X_i method, a value of 118 A was found. The $(X_{ox} + X_i)$ vs. $t_{ox}/(X_{ox} - X_i)$ plot obtained is shown in Fig. 5.6 and the rate constants found from the straight line fit were:

B = 28520 A^2/min

B/A = 8.66 A/min.

which are practically the same values used in the simulation (< 0.25% error). This analysis shows that the optimum X_i technique should be used for extracting the linear and parabolic rate constants.

Table 5.2 summarizes the results obtained for the rate constants when using other values of X_i between 100 A and 140 A as shown in Fig. 5.6. From Table 5.2 it is seen that in the case of X_i = 140 A the error in B and B/A is approximately 30% and 10%, respectively. In this case, X_i is only 22 A away from the optimum value of X_i.

Table 5.2

X_i (A)	$B(A^2/min)$	$B/A(A^2/min)$	% Error B	% Error B/A
100	24360	9.36	-14.6%	+ 7.97%
110	26030	9.01	- 8.75%	+ 3.95%
120	29410	8.56	+ 3.00%	- 1.17%
130	32970	8.19	+15.6%	- 5.46%
140	36640	7.88	+28.5%	- 9.04%

Therefore, using a value of X_i different from the optimum value would result in a good empirical fit to the data beyond a thickness X_c. However, the details of the physics of the oxidation process might be lost. Consequently, the rate constants should be obtained either by the optimum X_i technique or the inverse rate technique. These methods have been shown to yield the same results and still contain all the physical information of the problem involved.

In this investigation, <100>, <111>, and <110> oriented silicon wafers were oxidized in dry O_2 up to a thickness of 500 - 600 A. The growth was monitored by a high-temperature in-situ automated ellipsometer [5.9]. The linear and parabolic rate constants were obtained by optimizing X_i as detailed earlier. Table 5.3 summarizes the results obtained for 5 different temperatures in the 800 - 1000°C range. Besides values of B and B/A, the optimum value of X_i is also shown. It should be noted that for <100> and <111>, $X_{i(opt)}$ decreases with temperature. For the case of <110>, $X_{i(opt)}$ is practically constant with temperature.

5.2.4. Parabolic Rate Constant

The parabolic rate constants are shown for <100>, <111>, and <110> in Fig. 5.7a, 5.7b, and 5.7c, respectively. The values obtained by the Optimum

202

X_i method are plotted for 800°, 850°, 900°, 950° and 1000°C (open circles). For <100> and <111>, values for the linear rate constant are also included for 1200°C and 1100°C (solid circles). These values were obtained by applying the same technique to 1200°C and 1100°C data from the literature [5.9]. It is seen that the parabolic rate constant exhibits a breakpoint in activation energy between 950° and 1000°C.

Table 5.3

Rate Constants Using Opt X_i

T_{ox} (°C)	<100>	<111>	<110>
1000	28600	26600	14300
	8.65	16.4	25.4
	112	117	90
950	12100	10800	6590
	4.01	9.91	10.4
	105	105	98
900	5590	6500	4000
	2.08	4.60	4.61
	85	88	98
850	1890	2980	1840
	0.93	2.09	1.75
	83	80	100
800	660	1350	855
	0.43	0.90	0.72
	70	68	100

B (Å^2/min)

B/A (Å/min)

$X_{i(opt)}$ (Å)

Above this break temperature, the parabolic rate constant is independent of the substrate orientation and the process has the same activation energy of $(1.0 \pm .2)$ eV for both $\langle 100 \rangle$ and $\langle 111 \rangle$.

Below the break temperature, the parabolic rate constant is a function of the orientation of the underlying substrate. It exhibits a low-temperature activation energies of 2.22, 1.71 and 1.70 eV for $\langle 100 \rangle$, $\langle 111 \rangle$ and $\langle 110 \rangle$ respectively.

One possible interpretation of this behavior is the effect of the state of stress in the oxide on the diffusion process. Viscous flow is known to occur in SiO_2 films on silicon at a temperature around 960°C [5.12]. This is in the range where B breaks from its low-temperature orientation-dependent behavior to its high-temperature orientation-independent behavior. Below 960°C, the oxide has not flowed and is expected to be under stress. If the transport process depends on that stress, then the magnitude and the activation energy of B should be orientation-dependent, reflecting the dependence of the stress on the underlying substrate. It is observed that all three orientations exhibit different magnitudes of B. However, $\langle 111 \rangle$ and $\langle 110 \rangle$ have similar activation energies. Above 960°C, the silicon dioxide layer experiences viscous flow to relieve the stress. Transport then proceeds with a lower activation energy. The oxide bears no signature of the underlying substrate, hence, the magnitude and activation energy for the different orientations are identical.

5.2.5. Linear Rate Constant

The linear rate constants are shown for $\langle 100 \rangle$, $\langle 111 \rangle$ and $\langle 110 \rangle$ in Fig. 5.8a, 5.8b, and 5.8c respectively. The values obtained by the optimum X_i method are plotted for 800°, 850°, 900°, 950°, and 1000°C (open circles). For $\langle 100 \rangle$ and $\langle 111 \rangle$, values for the linear rate constant are also included for

1200°C and 1100°C (solid circles). These values were obtained by applying the same technique to 1200°C and 1100°C data [5.9] from the literature. It should be mentioned here that the error bars on the 1200 and 1100°C are rather large because the values of A involved are small, and any slight change in the straight line fit resulted in a small error in B, but in a large error in B/A. For example, the 1200°C data practically fitted a pure parabolic relationship. This mathematically, corresponds to a linear-parabolic relationship where A = 0, and B/A is infinite.

As in the case of the parabolic rate constant, the linear rate constant exhibits a break at approximately the same temperature. However, the higher temperature end exhibits a higher activation energy than the lower temperature end. Activation energies of 1.76, 1.74 and 2.10 eV are observed for ⟨100⟩,⟨111⟩ and ⟨110⟩ respectively. ⟨111⟩ and ⟨100⟩ have similar activation energies and the average ratio of ⟨111⟩ to ⟨100⟩ is 2.18. The high temperature activation energy is (2.75 ± 0.22) eV for ⟨100⟩ and ⟨111⟩.

These observations could also be interpreted by considering the effect of the stress in the oxide on the surface reaction. In the low-temperature regime, the silicon-silicon bond breaking energy is lowered by the presence of the stress in the oxide. In the high temperature range, the oxide has experienced viscous flow, the stress is therefore considerably reduced and the silicon-silicon bond breaking energy has to be supplied thermally. Hence the higher activation energy.

5.2.6. Empirical Fit to the Oxidation Rate Constants

For empirical fitting purposes, the linear and parabolic rate constants for ⟨100⟩, ⟨111⟩, and ⟨110⟩ were least square fitted to a single-activation energy process in the 800°C - 1000°C range. Table 5.4 contains pre-exponential constants (K) and activation energies (ΔE) for both B and B/A for ⟨100⟩, ⟨111⟩, and ⟨110⟩. Also shown are the percentages of error in the fit.

Table 5.4

	B (Å^2/min)	B/A (Å/min)	T_{ox} (°C)	% Error in B	% Error in B/A
<100>	$K = 1.73 \times 10^{13}$ Å^2/min $\Delta E = 2.22$ eV	$K = 7.35 \times 10^{7}$ Å/min $\Delta E = 1.76$ eV	1000	-0.8	5.8
			950	-3.9	-5.6
			900	8.7	-0.4
			850	-2.3	-3.4
			800	-1.2	+4.0
<111>	$K = 1.36 \times 10^{11}$ Å^2/min $\Delta E = 1.71$ eV	$K = 1.32 \times 10^{8}$ Å/min $\Delta E = 1.74$ eV	1000	11.5	-6.3
			950	-14.4	8.1
			900	2.5	1.4
			850	-0.3	-1.0
			800	2.5	-1.6
<110>	$K = 3.78 \times 10^{10}$ Å^2/min $\Delta E = 1.63$ eV	$K = 4.73 \times 10^{9}$ Å/min $\Delta E = 2.10$ eV	1000	6.3	6.4
			950	-10.3	-4.6
			900	4.9	-1.5
			850	-1.0	-5.9
			800	0.9	6.2

5.3 ANALYSIS OF THE OXIDATION PROCESS IN THE THIN-REGIME

The kinetics of SiO_2 growth were monitored in-situ at high temperature by an automated ellipsometer. The frequency with which data points were taken is very well suited to study the rapidly changing physical mechanisms involved in the oxidation in the thin regime.

It was observed tht the growth kinetics followed a linear-parabolic behavior for thicknesses of 350 Å and beyond. Below 350 Å, the oxidation rate is enhanced over what is predicted by the linear-parabolic expression. The analysis of this enhancement could be expressed by one of the following expressions:

$$\frac{dX_{ox}}{dt} = \frac{B}{2X_{ox} + A} + \Delta_e (X_{ox}, t) \qquad (5.8)$$

or

$$\frac{dX_{ox}}{dt} = \frac{B}{2X_{ox} + A} (1 + \Delta_r(X_{ox}, t)) \qquad (5.9)$$

where $\Delta_e (X_{ox}, t)$ or $\Delta_r(X_{ox}, t)$ are enhancement terms which decrease with increasing thickness and time. As time progresses and the oxide thickness increases, the linear parabolic behavior is approached and the enhancement terms $\Delta_e (X_{ox}, t)$ or $\Delta_r (X_{ox}, t)$ disappear. It should be noted that both these approaches are equivalent because, in the thin regime, the linear-parabolic term is approximately equal to the linear rate constant:

The analysis of the enhancement, whether expressed as an excess term (Δ_e) or a ratio term (Δ_r), could be done either in the time domain or the thickness domain. Both these approaches were used.

5.3.1 Analysis of the Thin Regime in the Thickness Domain

A typical plot of the excess oxidation rate as a function of thickness is shown in Fig. 5.9. This is the case of <100> oriented substrate for temperatures in the 800 - 1000 °C range. The excess term was fitted to 2 decaying thickness exponentials:

$$\frac{dX_{ox}}{dt} = \frac{B}{2X_{ox} + A} + C_1 e^{- \frac{X_{ox}}{L_1}} + C_2 e^{- \frac{X_{ox}}{L_2}} \qquad (5.10)$$

207

For the process described by (C_2, L_2) it was found that:

1. L_2 is independent of temperature in the 800-1000°C range. Its value is L_2 = 69 A, 78 A and 60 A for <100>, <111> and <110> respectively.

2. C_2 fits a single-activation-energy plot as shown in Fig. 5.10. The pre-exponential constants and the activation energies are shown in Table 5.5. It is seen that <100> and <111> have similar activation energies. The average ratio of C_1<111>/C_2<110> ⋅ 1.52.

Table 5.5

$$C_2 = C_{20} \, e^{-\frac{\Delta E_{C_2}}{kT}}$$

	C_{20} (A/min)	ΔE_{C_2} eV
<100>	6.57×10^{10}	2.37
<111>	5.87×10^{10}	2.32
<110>	5.37×10^{8}	1.80

For the process described by (C_1, L_1) it was observed that:

1. L_1 slowly increases from 7.7 A to 800°C to 12.4 A at 1000°C for <100>. It increases from 10.9 to 17.1 A for <111>. It was found to be a constant of 14.7 A independent of temperature for <110>.

2. C_1 is shown in Fig. 5.11 for all three orientations.

In using the obtained values of C_1, C_2, L_1 and L_2, B, B/A in equation 5.11 and integrating it numerically to fit the experimental data, it was found that the first decaying exponential $(C_1 e^{-X_{ox}/L_1})$ affected

208

the X_{ox} - t_{ox} relationship very slightly for thicknesses up to ≈ 50 Å. Therefore, neglecting this term resulted in numerical errors <5% in the fit to the experimental data. Hence, a good empirical expression for the oxidation rate could include the second decaying exponential term only in addition to the linear parabolic term. This is, in fact, the model used initially in SUPREM III.

5.3.2 Analysis of the Thin Regime in the Time Domain

The enhancement term was found to fit the following expression in the time domain:

$$\frac{dX_{ox}}{dt} = \frac{B + K_1 e^{-t/\tau_1} + K_2 e^{-t/\tau_2}}{2X_{ox} + A} \tag{5.12}$$

Plots of K_1, K_1/A, K_2, K_2/A, $1/\tau_1$, and $1/\tau_2$ vs. $1/T$ have exhibited well-behaved activated behavior. The activation energies in K_1 and K_1/A were found very close to those in B and B/A respectively. This suggests that the species involved in the first decaying time exponential are the same as those involved in B and B/A. One possible explanation is that in the initial stages of the oxidation the concentration of oxidants in the oxide is higher than what is predicted from solid solubility. However, the concentration then relaxes with a time constant τ_1 to the solid solubility value. The activation energy of K_2/A was found to be independent of orientation and is ≈ 2 eV. Analysis of the oxidation data at partial pressures <1 atm will help elucidate the different mechanisms possibly involved in the oxidation process, and may enable us to further refine the thin oxide model in SUPREM III.

REFERENCES

[5.1] B. E. Deal and A. S. Grove, "General Relationship for the Thermal Oxidation of Silicon," _J. Appl. Phys._, 36, Dec. 1965, 3770-3778.

[5.2] J. D. Meindl, et. al., "Final Report on Computer-Aided Engineering of Semiconductor Integrated Circuits," Stanford Electronics Laboratories, Stanford University, TRDXG569, July 1979.

[5.3] A. G. Revesz and R. J. Evans, "Kinetics and Mechanism of Thermal Oxidation of Silicon with Special Emphasis on Impurity Effects," _J. Phys. Chem. Solids_, 30, 1969, 551-564.

[5.4] S. W. Ing, R. E. Morrison and J. E. Sandor, _J. Electrochem. Soc._, 109, 1962, 221.

[5.5] E. A. Irene, "Silicon Oxidation Studies: Some Aspects of the Initial Oxidation Regime," _J. Electrochem. Soc._, 125, 1978, 1708-1714.

[5.6] G. F. Derbenwick and R. E. Anderson, private communication.

[5.7] Y. J. van der Meulen, "Kinetics of Thermal Growth of Ultra-Thin Layers of SiO_2 on Silicon. Part 1: Experiment," _J. Electrochem. Soc._, 119, 1972, 530.

[5.8] R. Ghez and Y. J. van der Meulen, "Kinetics of Thermal Growth of Ultra-Thin Layers of SiO_2 on Silicon. Part 2: Theory," _J. Electrochem. Soc._, 119, 1972, 1100.

[5.9] J. D. Plummer et. al., "Computer-Aided Design of Integrated Circuits Fabrication Processes for VLSI Devices," Stanford Electronics Laboratories, Stanford University, TRDXG501-81, July 1981.

[5.10] J. Blanc, "A Revised Model for the Oxidation of Si by Oxygen," _Appl. Phys. Lett._, 33, 1978, 424.

[5.11] D. W. Hess and B. E. Deal, "Kinetics of the Thermal Oxidation of Silicon in O_2/HCl Mixtures," _J. Electrochem. Soc._, 124-5, 1977, 735-739.

[5.12] E. P. Eernisse, "Viscous Flow of Thermal SiO_2," _Appl. Phys. Lett._, 30-6, 1977, p.290.
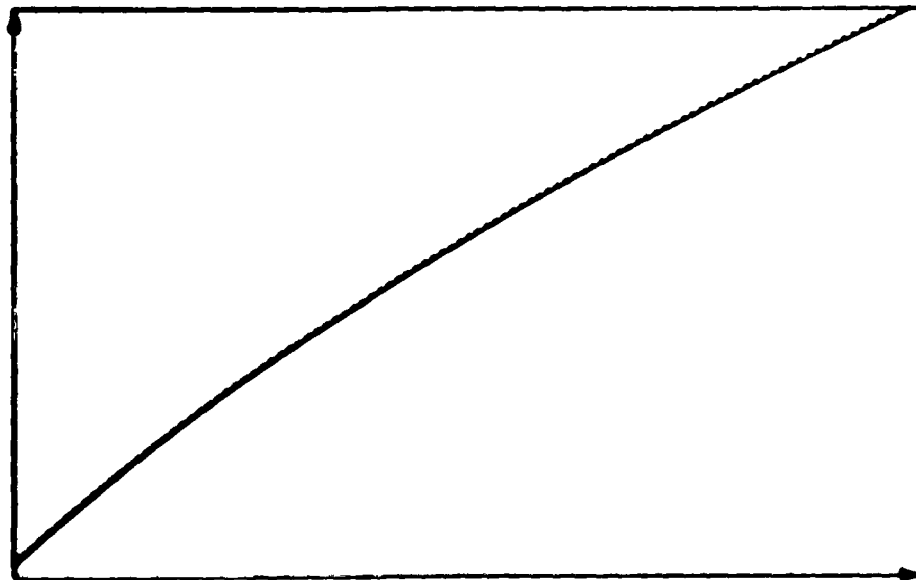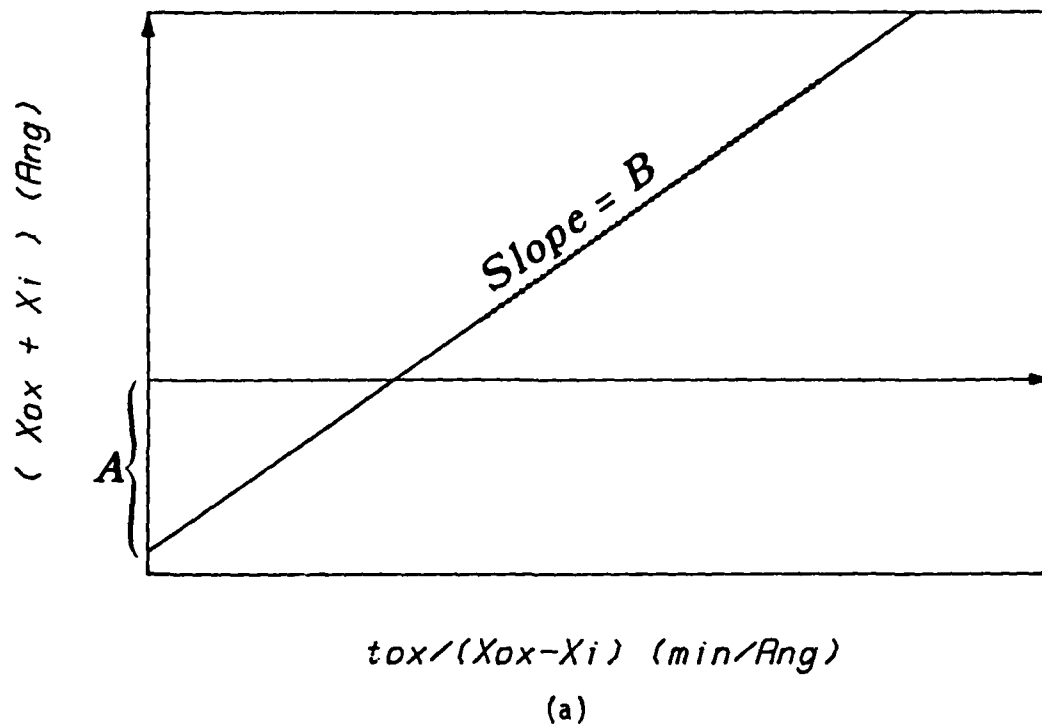
Dry Oxygen

(a)
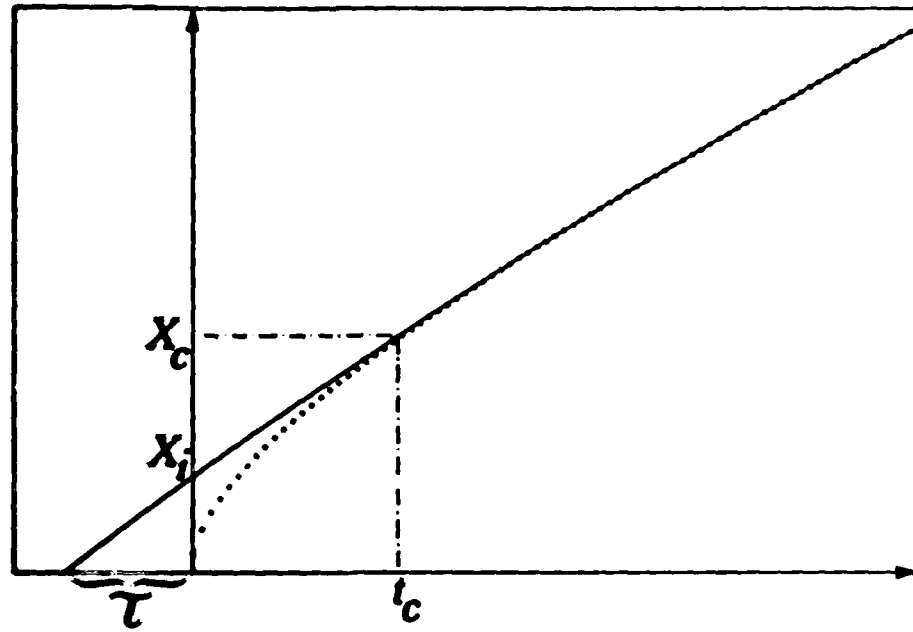
Wet Oxygen

Oxidation Time (min)

(b)

Fig. 5.1: Modeling oxide growth with a linear-parabolic relationship: (a) dry oxygen, (b) wet oxygen.

Rate Constants Extraction using X_i
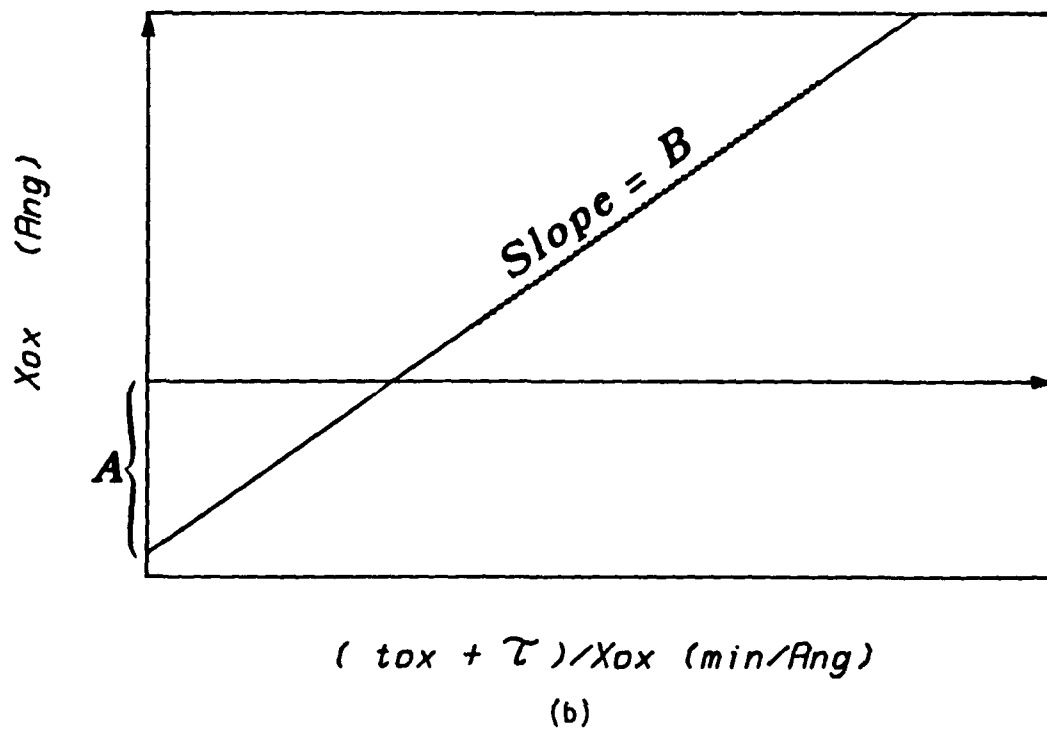

(a)

Rate Constants Extraction using τ

(b)

Fig. 5.2: Linear and parabolic rate constants extraction: (a) using X_i, (b) using τ.
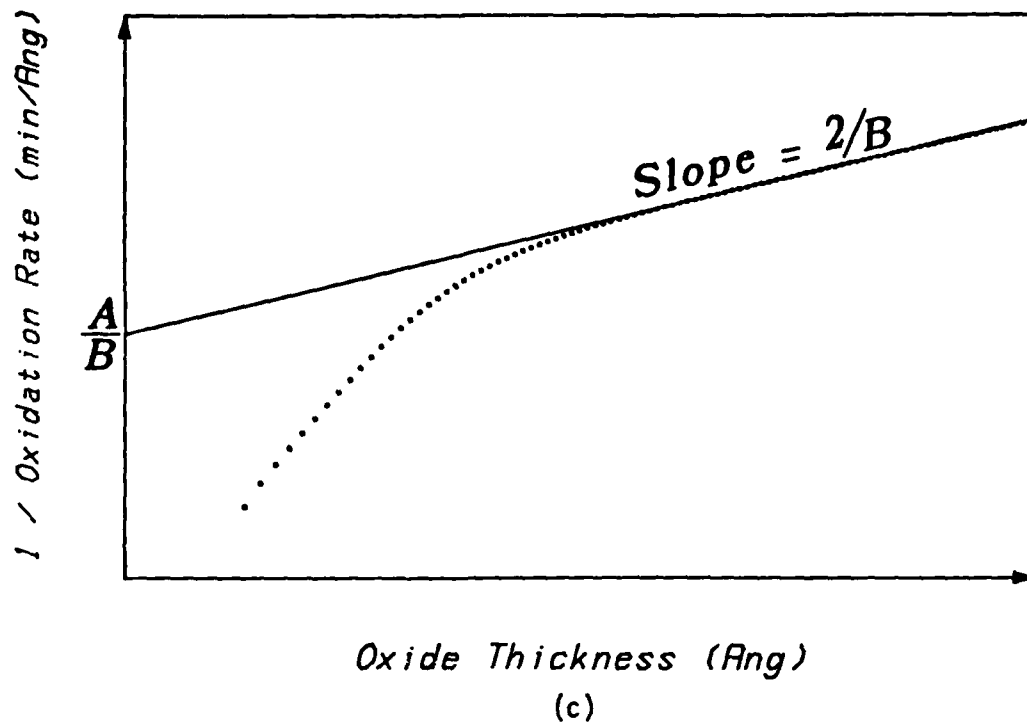
(c)

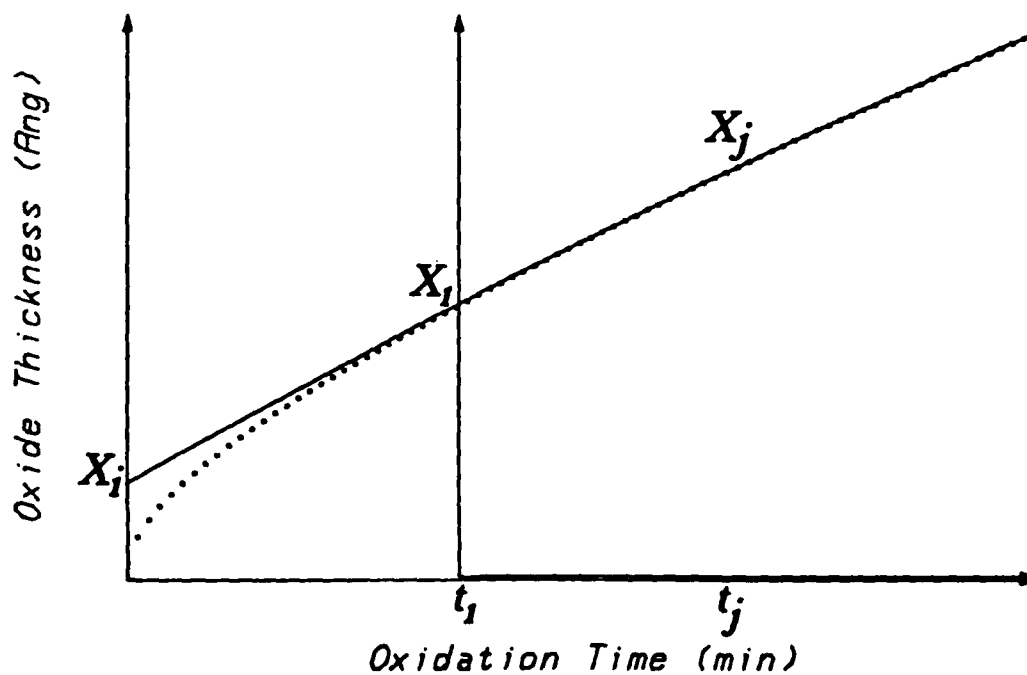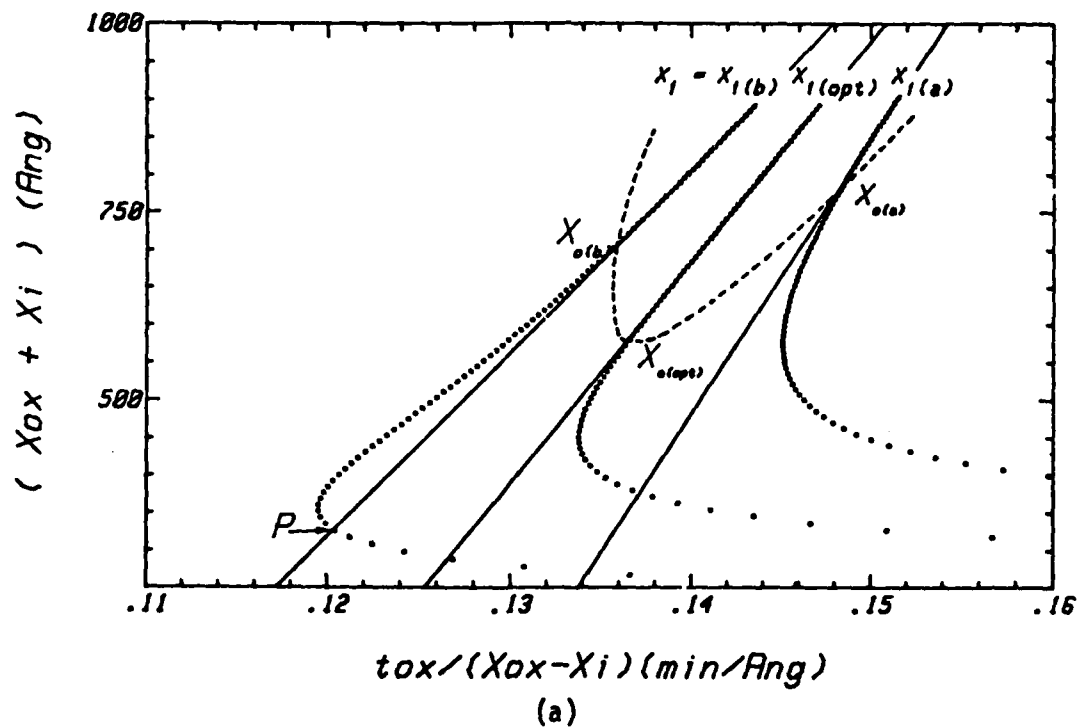Fig. 5.2c: Linear and parabolic rate constants, extraction using an inverse rate plot vs. thickness.

Fig. 5.3: Linear and parabolic rate constants extraction by shifting the time axis to t_1 and fitting the data beyond X_i.
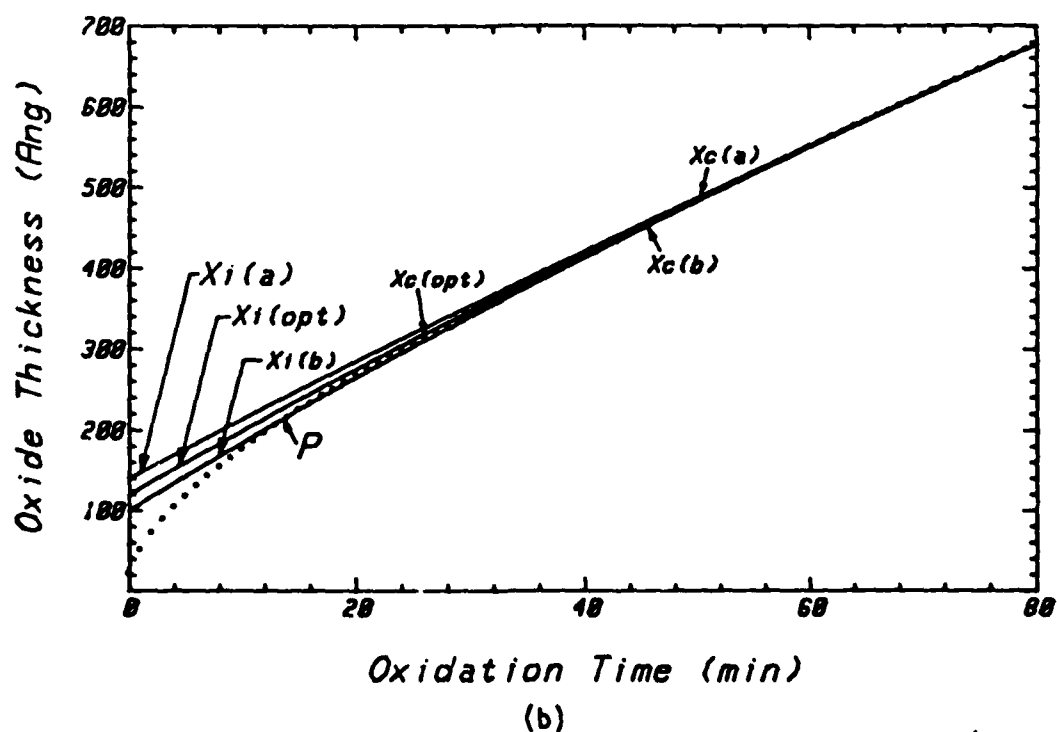
Fig. 5.4: Effect of the value of X_i on: (a) the extracted values of B and B/A, and (b) on the fit to the experimental data.
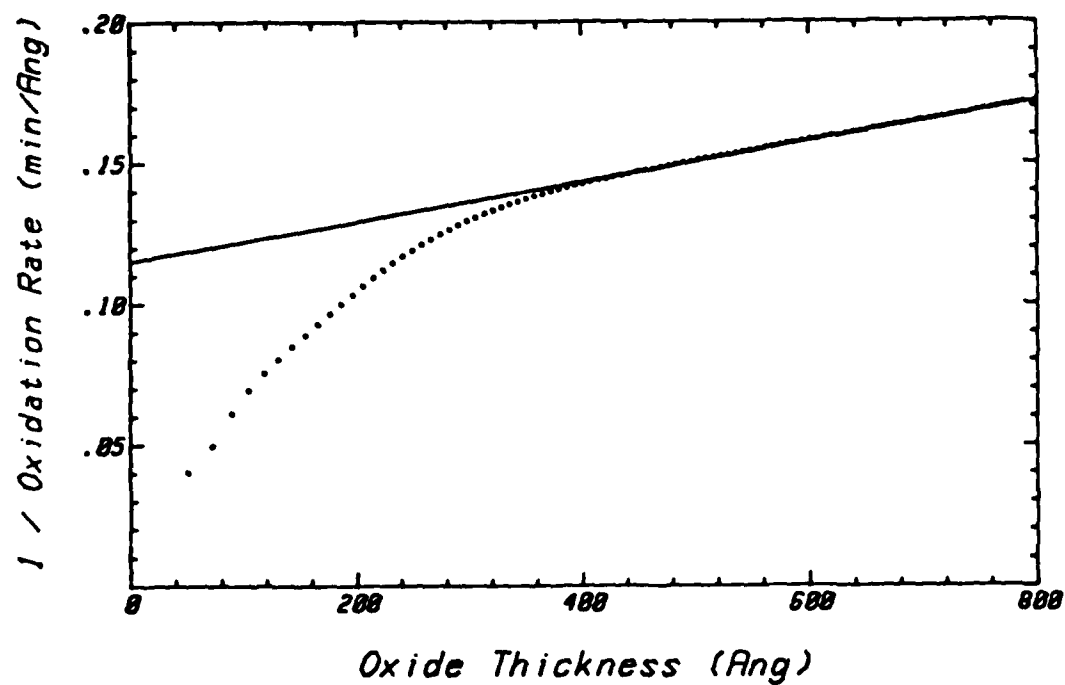
Fig. 5.5: Rate constants extraction using the inverse rate method.
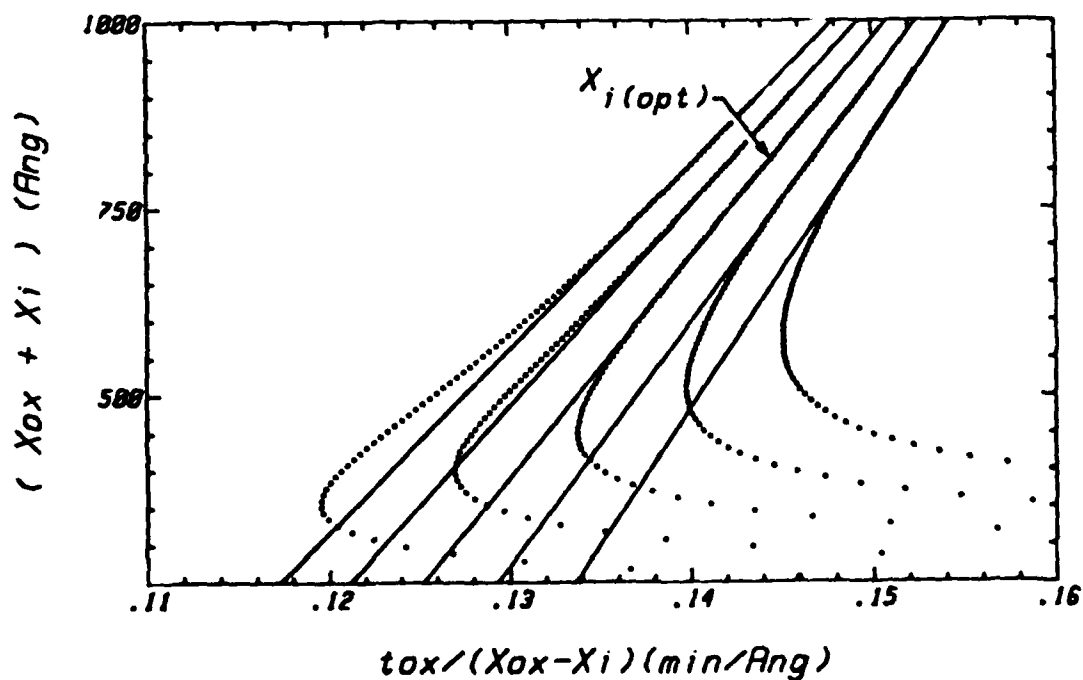
Fig. 5.6: Effect of X_i on extracted rate constants. The plots shown are for X_i = 100, 110, 120, 130 and 140 Å.

215
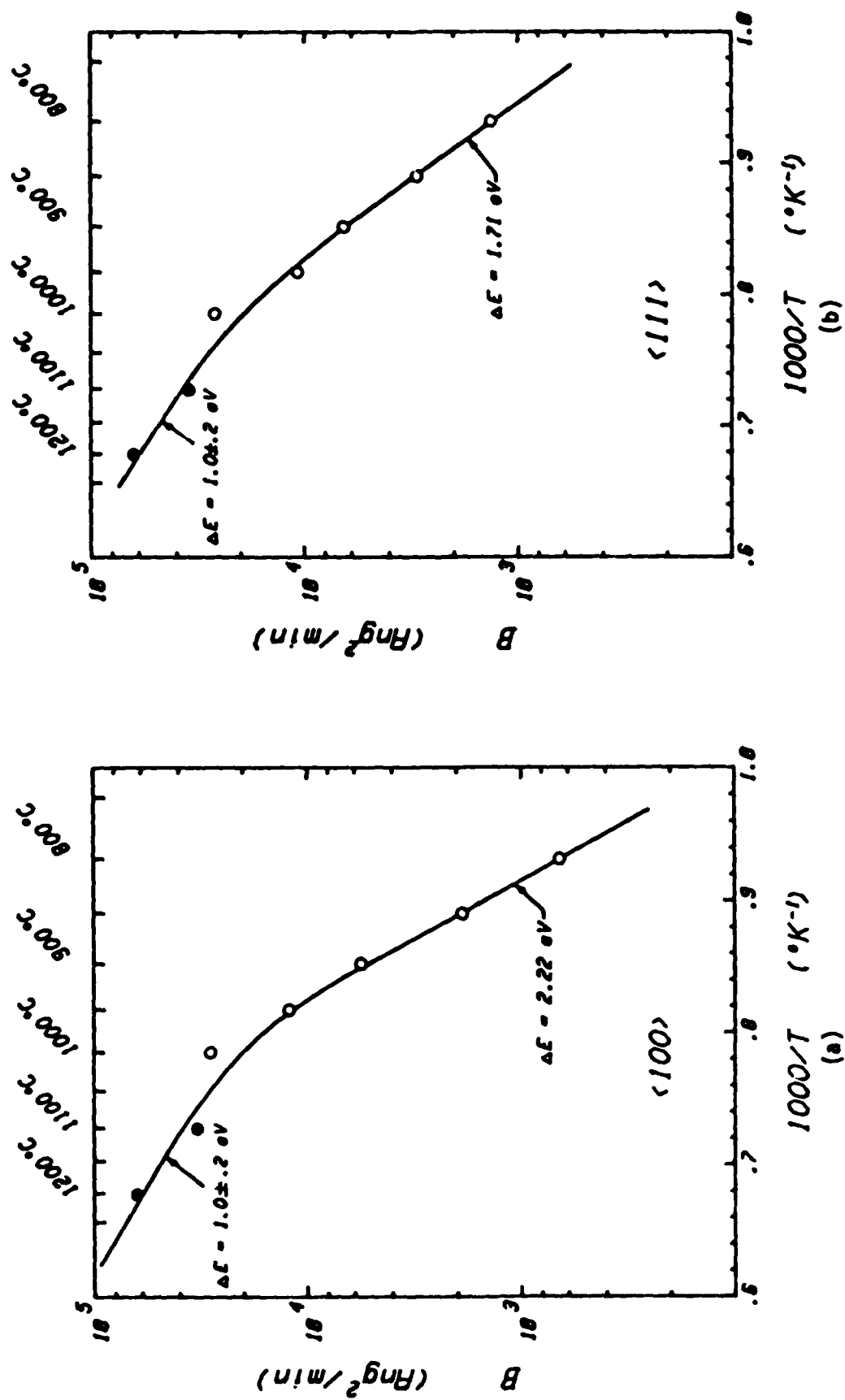
Fig. 5.7: Parabolic rate constant for the oxidation of: (a) <100>, (b) <111>, silicon in dry oxygen.
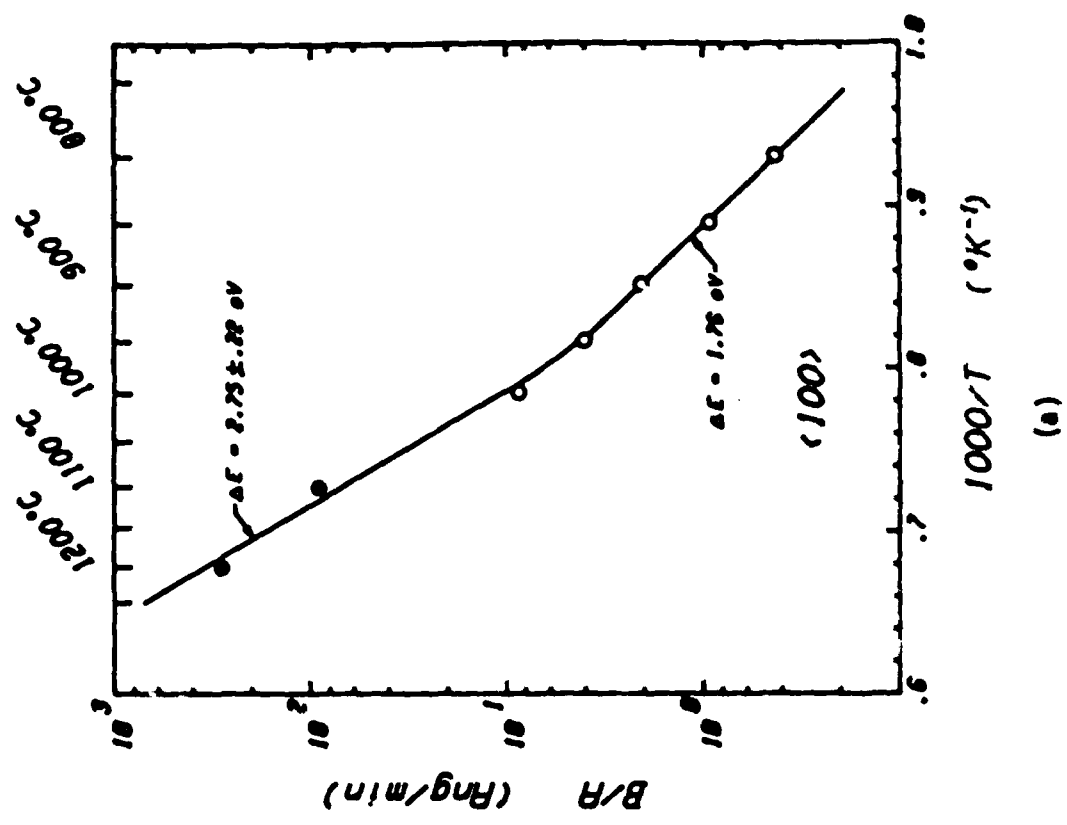
216

Fig. 5.8 Linear Rate Constant for the oxidation of (a)⟨100⟩ (b)⟨111⟩ and (c)⟨110⟩ Silicon in dry oxygen.

Fig. 5.7 Parabolic Rate Constant for the oxidation of (a)⟨100⟩ (b)⟨111⟩ and (c)⟨110⟩ Silicon in dry oxygen.

217
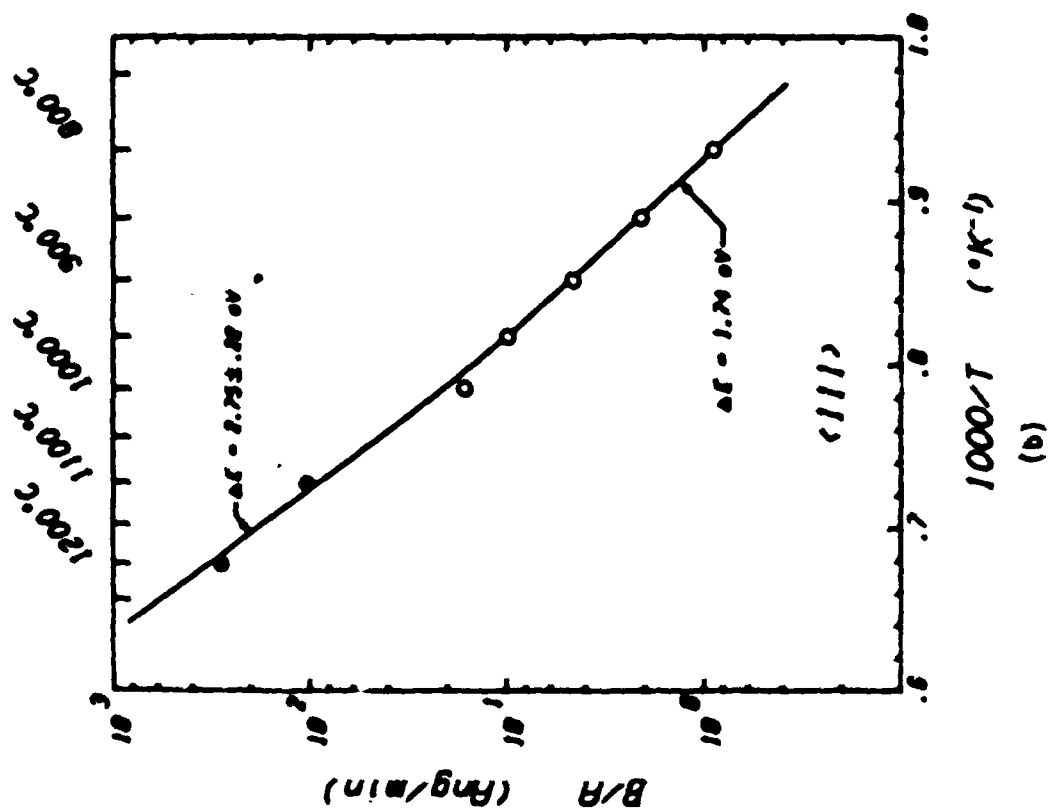
Fig. 5.8 Linear Rate Constant for the oxidation of (a)<100> (b)<111>and (c)<110> Silicon in dry oxygen.

Fig. 5.8 Linear Rate Constant for the oxidation of (a)<100> (b)<111>and (c)<110> Silicon in dry oxygen.
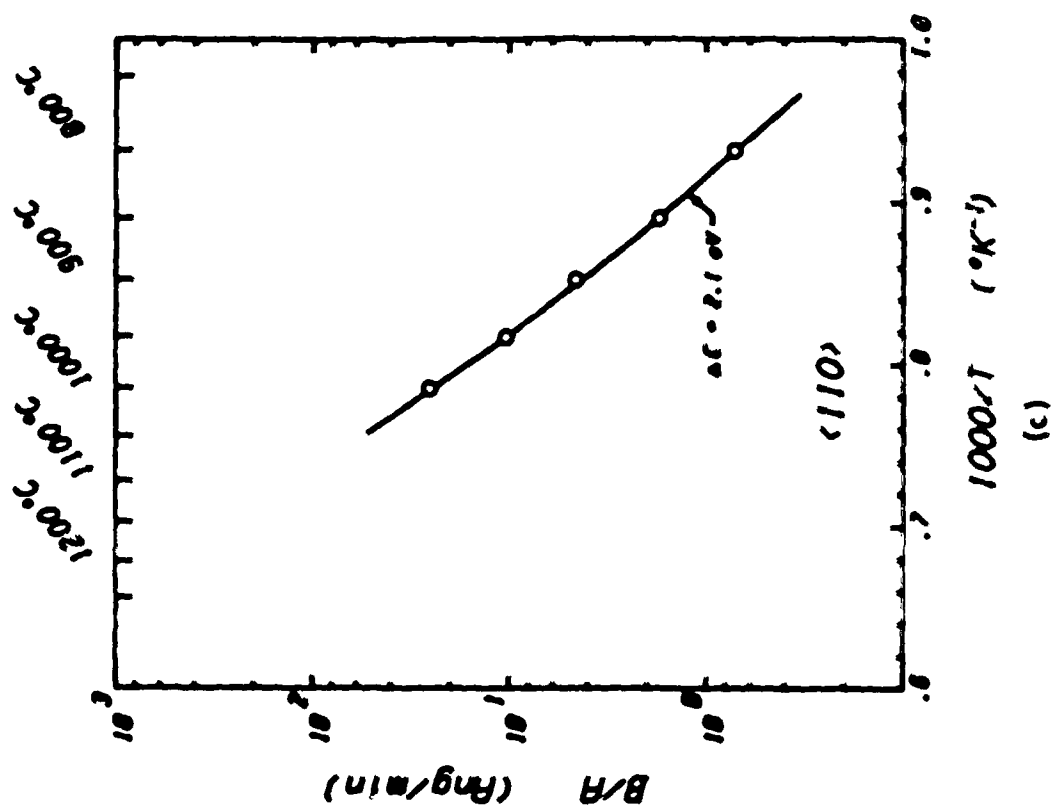
218

‹100›, Dry Oxygen, Lightly-doped Silicon
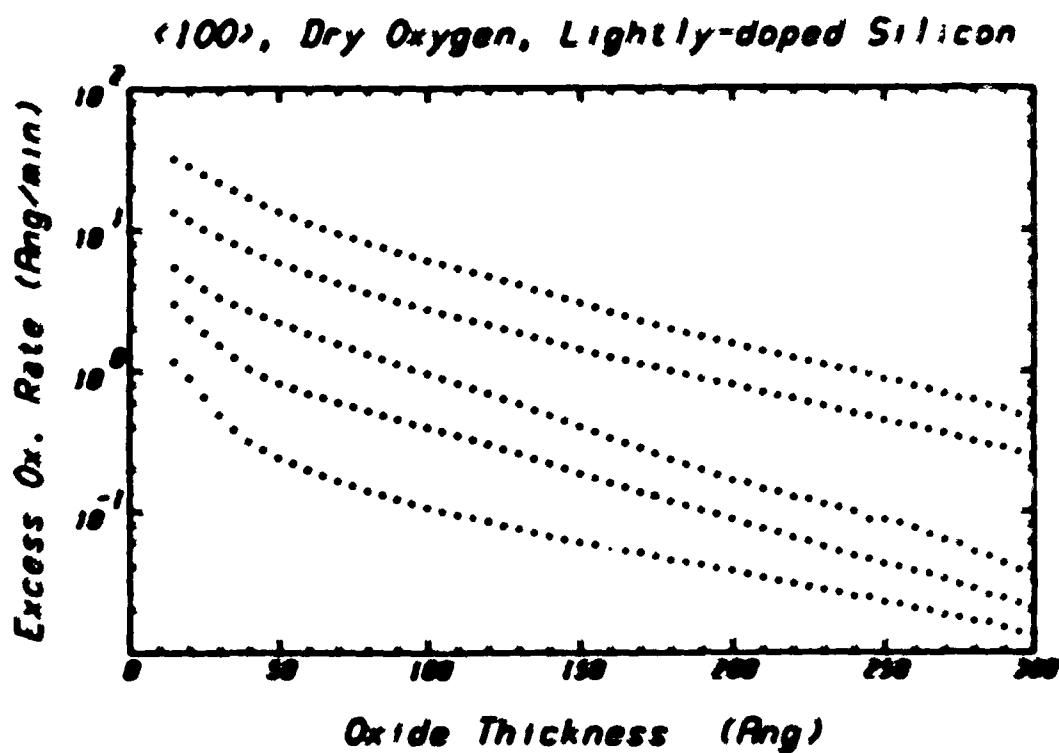
Fig. 5.9: Typical plot of the excess in the oxidation rate as a function of thickness for different temperatures in the 800°C - 1000°C range.
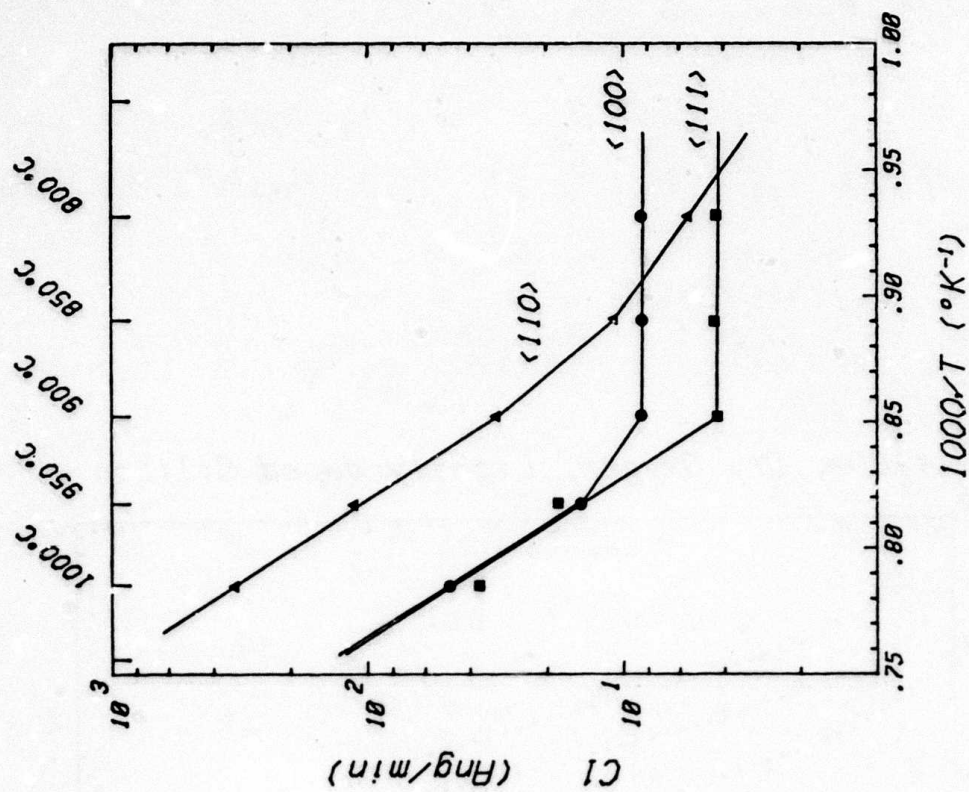
Fig. 5.11: Temperature dependence of the pre-exponential constant in the first exponentially decaying term for ⟨100⟩, ⟨111⟩ and ⟨110⟩.
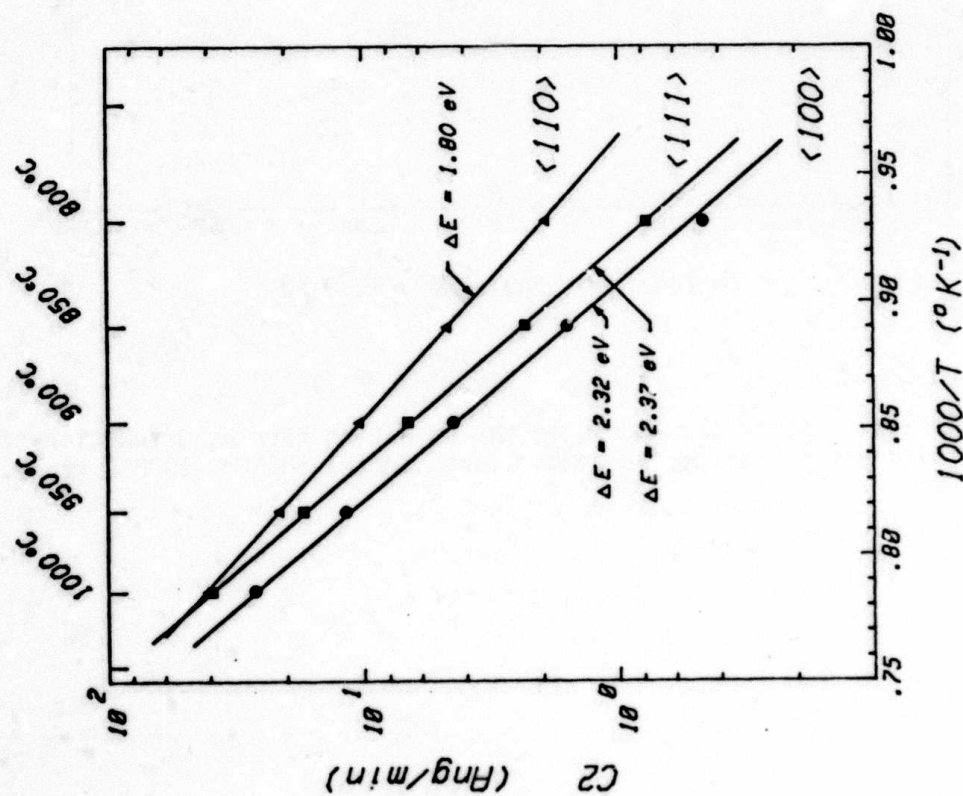
Fig. 5.10: Temperature dependence of the pre-exponential constant in the second decaying term for ⟨100⟩, ⟨111⟩ and ⟨110⟩.

220

6. THERMAL OXIDATION OF SILICIDES

K. Saraswat

6.1 INTRODUCTION

Continual advancements in integrated circuits technology are resulting in smaller device dimensions, increased chip area, and therefore an increased number of components on a chip, which require very complex interconnection wiring for communication within the chip. The smaller size of devices makes them faster, which should improve circuit performance. However, because of the larger chip size, the interconnection lines become very long, and the time delays associated with the parasitic capacitance and resistance can become appreciable. Thus even with very fast devices, the overall performance of a large circuit could be seriously affected by the RC delays of the interconnections [6.1]. Proper selection of material within the constraints of fabrication technology, can result in minimization of RC delay time. In a multilayer interconnection technology, the layers incorporated early in the processing are required to be able to withstand thermal oxidation in O_2 and H_2O. Polycrystalline silicon has been used in this situation, but its high resistivity results in very high RC delay time. Composites of polycrystalline silicon and silicides of Mo, W, Ti and Ta have been proposed as alternatives to polycrystalline silicon.

Several investigations have been made which studied the thermal oxidation of these silicides, and it has generally been found that silicides of W, Mo and Ti can be thermally oxidized in O_2 and H_2O ambients [6.2-6.5]. However, considerable controversy exists about the kinetics of oxidation of $TaSi_2$. In the case of $TaSi_2$, Murarka, et. al. [6.6] observed that the oxidation of $TaSi_2$/poly-Si composite structure occurs readily in H_2O, but is negligible in an O_2 ambient.

Our prior work [6.7], and that of Razouk, et. al. [6.8] to the contrary, indicate that oxidation occurs in both ambients. In this work we have attempted to obtain a better understanding of the oxidation behavior of $TaSi_2$. The overall goal is a a set of oxidation models for all these silicide structures suitable for incorporation in SUPREM III.

6.2 EXPERIMENTAL PROCEDURE

The tantalum silicide films were deposited by simultaneous rf diode sputtering of tantalum and silicon from high-purity targets. The details of the technique are described in the literature [6.9]. The data presented here are for 0.3 μm thick silicide films, deposited on high resistivity <111> oriented silicon and oxidized silicon wafers. Auger electron spectroscopy and electron microprobe analysis indicated that in the as-depostied films, the atomic ratio between silicon and tantalum was approximately 2:1. X-ray diffraction analysis showed no specific peaks of Ta, Si or $TaSi_2$, indicating that the as-deposited films were amorphous. Therefore, prior to the oxidation, the samples were annealed at the oxidation temperature (1000°C or 1100°C) in argon for 10 minutes to homogenize and crystallize the structure of the tantalum silicide. The annealing resulted in polycrystalline films of $TaSi_2$ free of any preferred orientation. Following the anneal, the samples were oxidized according to three different conditions: (1) 60 min. in dry O_2 at 1000°C, (2) 60 minutes in dry O_2 at 1100°C, and (3) 10 minutes in steam (Ar bubbled through 95°C H_2O) at 1000°C. Following the oxidation, the wafers were annealed for 10 min. in Ar and slowly pulled out of the furnace. The oxidized samples were examined by auger electron spectroscopy, X-ray diffraction and scanning electron microscopy.

6.3 RESULTS AND DISCUSSION

In all cases, i.e., in O_2 and H_2O ambients and on $TaSi_2/Si$ and $TaSi_2/SiO_2$ structures thermal oxidation was observed. The scanning electron micrographs

222

showed that the resulting surface was very rough for the oxidation of $TaSi_2/SiO_2$ samples and in some cases, cracks and lifting of the films from the substrate were observed. The surface was considerably smoother in the case of $TaSi_2/Si$ oxidation and no cracks or lifting were observed.

All the samples were analyzed by the X-ray diffraction technique where the film texture was obtained by measuring the diffracted intensity at twice the angle of incidence of an X-ray beam from 20 keV copper radiation. The results are shown in Figure 6.1 for all oxidation conditions. In the case of $TaSi_2/Si$ samples, only $TaSi_2$ diffraction peaks were observed, whereas for $TaSi_2/SiO_2$ samples, additional diffraction peaks of Ta_2O_5 were also observed. No other phase of tantalum silicide was observed. By comparing Fig. 6.1(c) and 6.1(f) it can be deduced that at the higher oxidation temperature of 1100°C, the intensity of $TaSi_2$ diffraction peaks reduced and those of Ta_2O_5 increased. The diffraction pattern of the annealed but not yet oxidized samples was very similar to those shown in Figs. 6.1(a), (b) and (c) and is therefore not reproduced here. No Ta_2O_5 formation could be detected in the annealed samples.

The samples were further examined by the auger sputter profiling technique and the results are shown in Figs. 6.2 and 6.3. It can be seen in both the figures that the bulk of the oxidized layers contain only silicon and oxygen and within the resolution of the auger spectrometer, no tantalum could be measured in the grown films. By comparing these auger profiles to those of known SiO_2 thermally grown on silicon, it was confirmed that the grown films are indeed Stoichiometric SiO_2. It is evident that $TaSi_2$ can be thermally oxidized in dry oxygen with and without the presence of free silicon underneath. Considerable broadening of the $TaSi_2$ and $TaSi_2/SiO_2$ interface can be observed in Figs. 6.2 and 6.3 and it increases with the thickness of the grown SiO_2. This broadening could be partly attributed to artifacts of the measurement technique, but partly also to interface roughening [8].

223

By combining the results of X-ray diffraction and auger sputter profiling, certain conclusions can be drawn regarding the mechanism of oxidation. For the oxidation of $TaSi_2$ deposited on silicon (Fig. 6.2), as the oxidation proceeds, the entire layer of $TaSi_2$ is displaced towards the substrate without any change in its thickness. From the X-ray diffraction results, we concluded earlier that no change in the crystal structure of the $TaSi_2$ was observed. This indicates that as long as free silicon is available underneath the silicide, tantalum is not oxidized at all. At the temperature of oxidation, Si is highly mobile in other silicides [6.2-6.5] and it is reasonable to assume that it diffuses rapidly in $TaSi_2$ also. Thus as the oxidation proceeds, silicon diffuses through the $TaSi_2$ and is oxidized by O_2 or H_2O. The composition of the $TaSi_2$ remains unchanged.

The situation is more complex if free silicon is not available. From the auger sputter profiling (Fig. 6.3) it can be seen that a layer of SiO_2 is still grown on top of the silicide, however, the X-ray diffraction analysis (Fig. 6.1) shows that formation of Ta_2O_5 is accompanied with that of SiO_2. Since no free silicon is available for oxidation, it is reasonable to assume that tantalum silicide is oxidized into Ta_2O_5 and SiO_2

$$4 \ TaSi_2 + 13 \ O_2 \rightarrow 2 \ Ta_2O_5 + 8 \ SiO_2 \qquad (6.1)$$
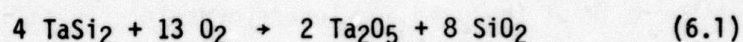
A similar reaction should take place with H_2O. Similar reactions have been observed in the oxidation of WSi_2 [6.2,6.3], $MoSi_2$ [6.10], and $TiSi_2$ [6.5] deposited on SiO_2. Since no tantalum was detected in the bulk of the grown SiO_2, the Ta_2O_5 probably remains at the interface or right next to it in the silicide layer. As the oxidation proceeds, more Ta_2O_5 is formed as the X-ray diffraction results show (Fig. 6.1).

In general, these results are similar to the oxidation of other silicides [6.2-6.5] but differ from the observation of Murarka [6.6] where no oxidation of $TaSi_2/Si$ was observed in dry oxygen.

6.4 SUMMARY

In summary, we have shown that $TaSi_2$ can be thermally oxidized in dry O_2 and steam at 1000°C and 1100°C. If free silicon is present underneath the $TaSi_2$, it diffuses through the silicide and is oxidized to produce SiO_2. In the other case, where no free silicon is present, $TaSi_2$ is directly oxidized to form SiO_2 and Ta_2O_5. This data will be useful in combination with our earlier work on WSi_2 and $MOSi_2$, in developing general kinetic models for SUPREM. The multilayer capability of SUPREM III requires a large number of new material models; this data will help to provide some of these models.

REFERENCES

[6.1] K. C. Saraswat and F. Mohammadi, IEEE Trans. Electron Dev., ED-29, 645 (1982).

[6.2] S. Zirinski, W. Hammer, F. d'Heurle, and J. Baglin, Appl. Phys. Lett., 33(1), 76 (1978).

[6.3] F. Mohammadi, K. C. Saraswat, and J. D. Meindl, Appl. Phys. Lett., 35(7), 529 (1979).

[6.4] T. Inoue and K. Koike, Appl. Phys. Lett., 33(9), 826 (1978).

[6.5] J.- R. Chen, M.- P. Houng, S.- K. Hsiung, and Y.- C. Liu, Appl. Phys., Lett., 37, 824 (1980).

[6.6] S. P. Murarka, D. B. Fraser, W. S. Lindenberger, and A. K. Sinha, J. Appl. Phys., 51, 3241 (1980).

[6.7] K. C. Saraswat, R. S. Nowicki and J. F. Moulder, Abstract from the Technical Program of the Electronics Materials Conference, Santa Barbara, CA, June 24-26, 1981.

[6.8] R. R. Razouk, M. E. Thomas, and S. L. Pressacco, Appl. Phys. Lett., (To be published).

[6.9] R. S. Nowicki and J. F. Moulder, J. Electrochem. Soc., 128, 562 (1981).

[6.10] J. B. Berkowitz - Mattuck, and R. K. Dils, J. Electrochem. Soc., 112, 583 (1965).
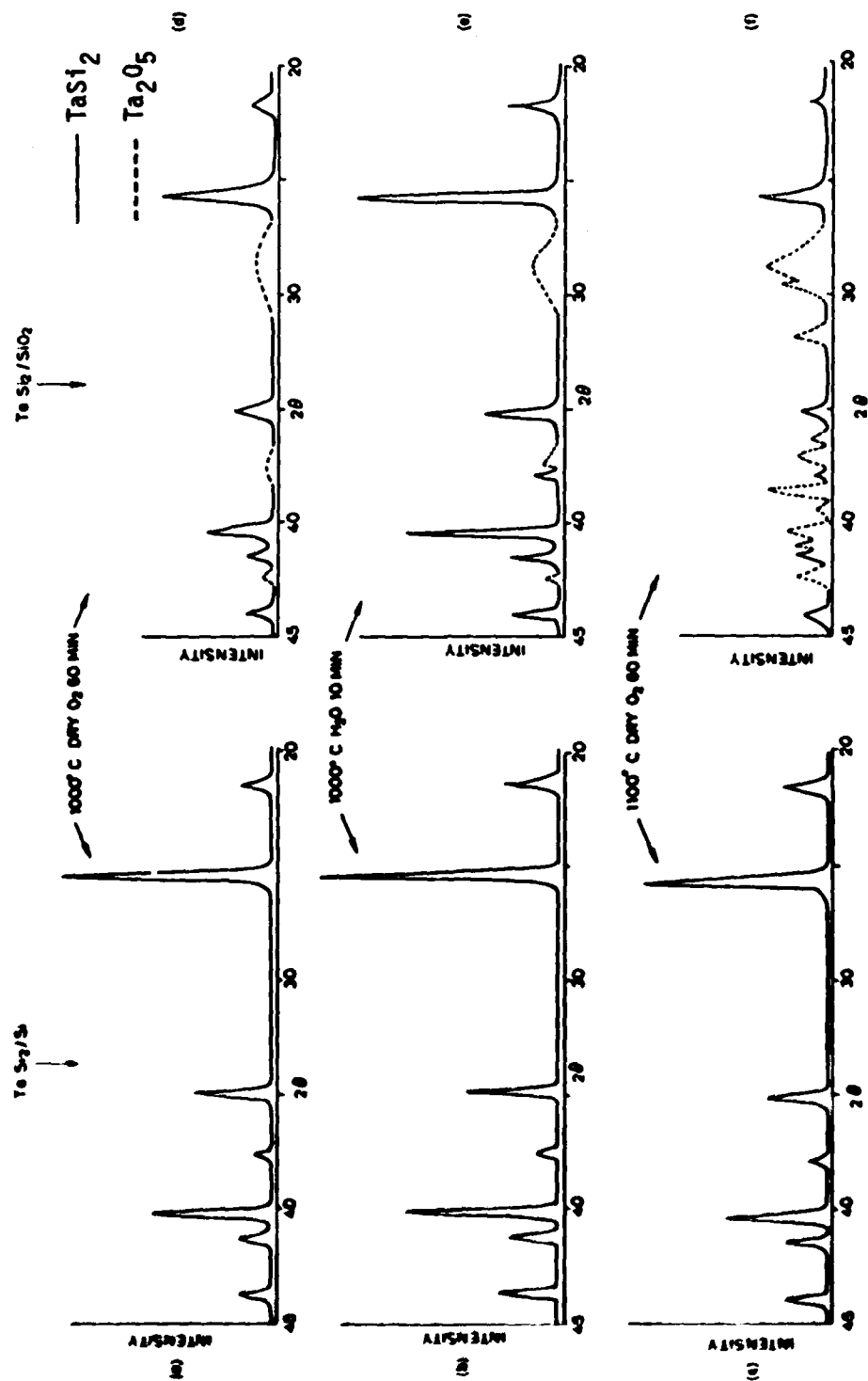
Fig. 6.1: X-ray diffraction pattern of the oxidized tantalum silicide films.
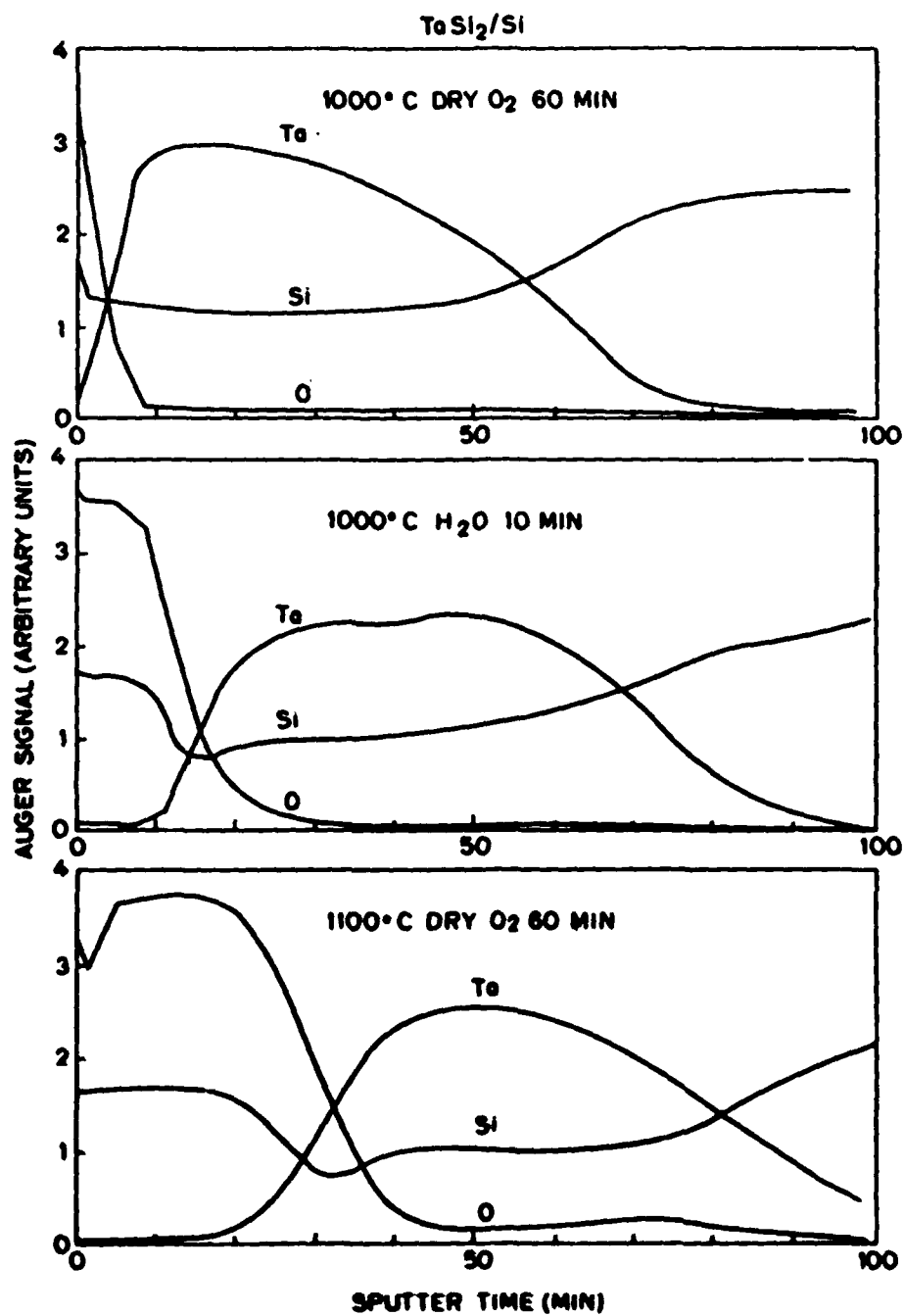
Fig. 6.2: Auger sputter profiles of tantalum silicide films deposited on <111> silicon and oxidized at (a) 1000°C in dry O_2 for 60 min., (b) 1000°C in H_2O for 10 min., and (c) 1100°C in dry O_2 for 60 min.

228
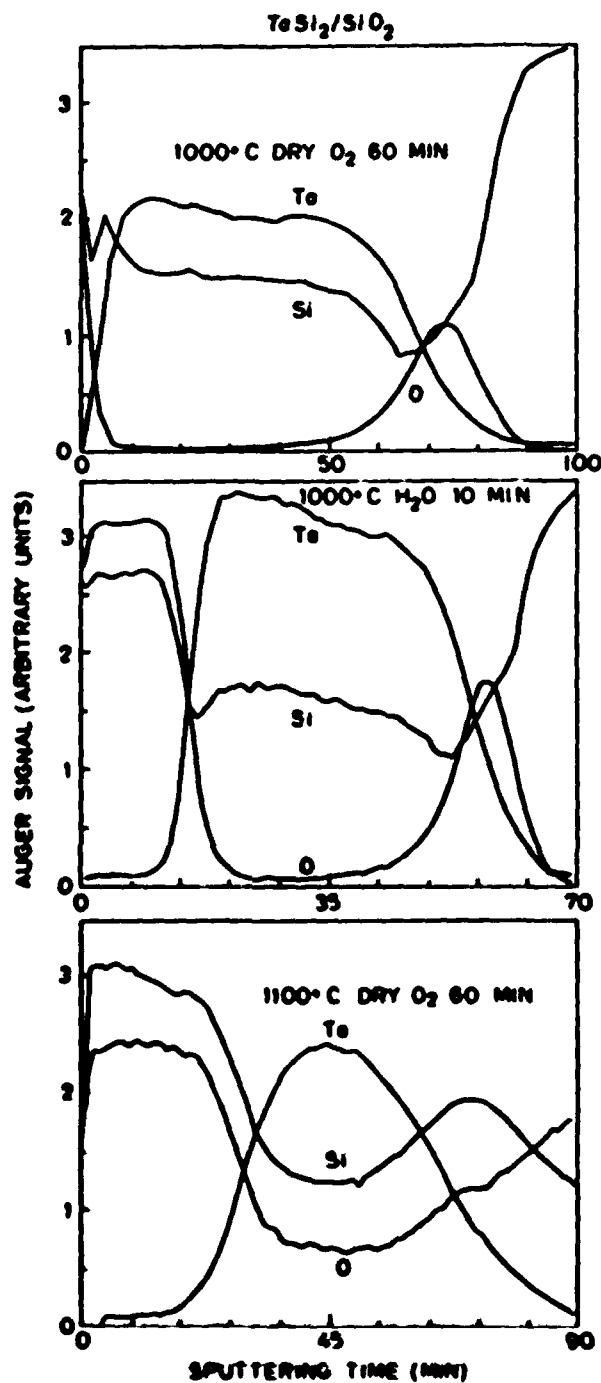
Fig. 6.3: Auger sputter profiles of tantalum silicide films deposited on oxidized silicon at (a) 1000°C in dry O_2 for 60 min., (b) 1000°C in H_2O for 10 min. and (c) 1100°C in dry O_2 for 60 min.

7.   Si/SiO$_2$ INTERFACE CHARGES

T. Akinwande, C. Ho, J. Plummer

7.1   INTRODUCTION

As device dimensions shrink, and with the increase in importance of oxide isolation, an improved understanding of the charges associated with the Si-SiO$_2$ interface is essential to the fabrication of IGFETs. Of particular interest are the fixed oxide charge density, N_f, and the interface state trap density, D_{it}, which affect the threshold voltage, I-V characteristics, leakage between devices, breakdown voltages and transfer efficiency in CCD's. These charges become more important in thin gate oxide IGFET's, tunnel MIS diodes and MESFETs with tunnel barriers as these are some of the devices being considered for VLSI. It is well established that these charges are dependent on processing conditions such as oxidation time and temperature, pull time, annealing ambient, temperature and duration, and material properties such as surface orientation. Previous studies have resulted in much valuable qualitative knowledge of the Si-SiO$_2$ interface and reasonable control of the associated charges. Nevertheless it is the concensus among interface specialists that numerous questions are yet to be answered. An example is the fact that there is no quantitative model that describes the dependence of these charges on process parameters and material properties. Such a model would be of great use in a program such as SUPREM, and moreover, it would be of great help in understanding the oxidation process as well as the optimization of devices and process technologies.

7.2   CURRENT UNDERSTANDING OF OXIDE CHARGES

The current understanding of the as-grown and annealed values of the fixed oxide charge density N_f is best illustrated by the Deal Q_f triangle [7.1] in Fig. 7.1. The upper curves represent as-grown values while the horizontal bottom curve represents the annealed values. Lamb et. al. [7.2] obtained a slightly

different behavior of the unnealed N_f values as shown in Fig. 7.2. The interface state trap density D_{it} follows the same pattern as the fixed oxide charge density N_f but D_{it} magnitudes can be further reduced by a low temperature hydrogen anneal [7.3] as in Fig. 7.3. These results have led to the development of several qualitative models to explain the origin of the charges. The most important models are discussed below.

7.2.1 Excess Si or Si_I model [7.4 - 7.8]

The annealed properties of the fixed oxide charge N_f have led to the belief that silicon point defects (Si_I) or incompletely oxidized Si atoms are the origin of N_f. The fixed charges have been correlated with oxidation rate [7.8] and thus an anneal in an inert ambient is equivalent to zero oxidation rate. Similarly the low temperature anneal properties of D_{it} in hydrogen ambients [7.3] and high temperature anneals in chlorine [7.5] have led to the belief that the origin of D_{it} is the small number of Si dangling bonds produced during the oxidation process that remain in electrical communication with the substrate. Closely related to the excess Si model is the oxygen vacancy model [7.6] in which a chemical reduction of the oxides by CO resulted in increased charge densities. There are also models which postulate that an SiO_x layer is responsible for the charges [7.7] and that annealing at high temperature is equivalent to driving the oxidation reaction to completion. Unfortunately chemical profiling techniques such as XPS, AES, ASP, RBS, SIMS, etc. are not sensitive enough to draw definite conclusions about the presence or absence of excess Si near the Si/SiO_2 interface.

7.2.2 Lattice Mismatch (strain) Models [7.4, 7.9-11]

Using lattice mismatch and free volume concepts, it has been shown that a significant amount of strain energy is stored in the bonds at the $Si-SiO_2$

231

interface [7.4]. It has been postulated that strained bonds and dangling bonds at the interface could lead to states within and outside the Si bandgap [7.9]. Similar to this is the Strain Induced Charge Transfer (SICT) model in which strain in the Si-O-Si bridging bonds leads to charge transfer between Si and O atoms [7.11]. Also similar to above mentioned models, is the idea that postulates that strain at the interface leads to a semiconductor layer with band tails. It has been suggested that these band tail states are the origin of interface trap states [7.10].

Other models include those in which SiO_2 and Si are believed to form a heterojunction. It is believed that a defect level exists in the narrow depletion layer formed between Si and SiO_2, allowing electrons at the surface to tunnel to various energy levels [7.12]. The Variable Energy VE model is similar in concept [7.13].

7.3 EXPERIMENTAL RESULTS AND DISCUSSION

It can be seen from the above discussion that much is still to be understood about the nature, origin and control of these charges. It was with this in mind that we began a series of experiments in the past year, to understand their properties and origins.

Experiments have been conducted on <111> and <100> oriented n-type Si

(a) to investigate the possible dependence of the annealed values of the interface charges on temperature,

(b) to investigate the behavior of the charges under thermal cycling and hence help determine if the charges have an equilibrium value related to temperature,

(c) to determine accurately the as-grown values of the oxide charges at the Si-SiO_2 interface as a function of temperature.

The variation of Ar annealed values of the fixed oxide charge N_f with temperature is shown in Fig. 7.4. Contrary to published results [7.1], we found the

232

fixed oxide charge density N_f decreased with an increase in temperature. The N_f values below 950°C were found to increase with a decrease in temperature, but above 950°C, N_f values were found to be almost constant. This is of considerable importance since coincidentally SiO_2 is known to become viscous at a temperature of ≈965°C. This may support the theory that a part of the charges at the interface are stress related. The value and scatter of N_f at 1200°C is rather high, and this may be due either to the reaction between Si and SiO_2 to form gaseous SiO [7.7], or to cooling transients. The same trend is observed for the D_{it} values as shown in Fig. 7.5, but a breakpoint is not quite as obvious in this case.

The results of thermal cycling experiments are shown in Fig. 7.5 for N_f. As can be seen from the plots, N_f values on ⟨111⟩ oriented Si do thermally cycle up and down as the Ar anneal temperature is changed, while no significant conclusions could be drawn on the ⟨100⟩ oriented wafers, because the values are within the noise range of the measurement system.

We may conclude from these results above that contrary to the Q_f triangle, the lower part of the triangle is likely not flat and there seems to be an equilibrium value of N_f and D_{it} when they are annealed in Argon which is closely related to the temperature of anneal, surface orientation, and the nature of the SiO_2 layer. The results of the ⟨111⟩ N_f thermal cycling experiment leads us to believe that N_f might be related to a thermodynamic state of the Si-SiO_2 interface, although more data is needed to solidify this conclusion.

The as-grown values of the oxide charges N_f and D_{it} are shown in Figs. 7.7 and 7.8 for both ⟨111⟩ and ⟨100⟩ orientations. There is a break point at 950°C which has not been previously observed. The higher slope of the ⟨100⟩ N_f and D_{it} values seem to contradict the belief that the N_f and D_{it} values of ⟨111⟩ oriented Si are 1.7 times those of the ⟨100⟩ orientation. There is a marked

233

difference in the N_f and D_{it} values of the two orientations at higher temperatures where oxidation is diffusion limited and oxidation rate is expected to be about equal on the two orientations. At the lower temperatures, the charge values are closer in magnitude and at these temperature, it is expected that oxidation rate should be surface orientation dependent. While we cannot draw any definite conclusions, it seems that apart from the dependence of the charges on oxidation rate, there is an activated process present at the interface. This has led us to suggest the dependence given below:

$$N_f - N_{fe} = K_1 \left(\frac{dx}{dt}\right)^n e^{E_A/kT} \tag{7.1}$$

where

$$N_f = \text{As grown value of fixed oxide charge}$$

$$N_{fe} = \text{Annealed value of fixed oxide charge}$$

$$\frac{dx}{dt} = \text{oxidation rate}$$

$$E_A = \text{activation energy}$$

$$T = \text{absolute temperature}$$

$$n = \text{power factor}$$

We do not have enough data to be able to confirm this relationship at the present time without further experiments.

Subsequent to the above results, it is imperative that we conduct further experiments to enable us to arrive at a quantitative model for the oxide charges. Obviously the issues involved are (a) surface orientation of the underlying silicon substrate, (b) type of oxidation, (c) oxidation rate, and (d) temperature. It is assumed that just like other oxidation related processes such as OISF and OED, [7.14,7.15] the charges are related to silicon point defects and hence

234

$$Si_I \propto \left(\frac{dx}{dt}\right)^n \tag{7.2}$$

The exponential factor in (7.1) may represent the details of the oxidation process at the Si/SiO_2 interface [7.16].

While a knowledge of the as-grown value of the charges is essential for understanding their generation, the most important quantity is the final annealed value of the charges. The most important parameters for anneal are surface orientation, anneal ambient, temperature and period of anneal. It has been suggested that for deposited oxides, the densification kinetics are governed by the expression [7.17]

$$D_{it} = D_{it\infty} + (D_{ito} - D_{it\infty}) e^{-t/\tau} \tag{7.3}$$

where $\tau = K_1 e^{-E_{A_1}/kT}$, time constant

$D_{it\infty} = K_2 e^{-E_{A_2}/kT}$, steady state value

E_{A_1}, E_{A_2} = activation energies

T = absolute temperature

It is our belief that the anneal kinetics of thermally grown oxide charges should have the same form as the kinetics of densification of deposited oxides. We are conducting experiments on the anneal kinetics of thermally grown oxides. We hope that these experiments will aid us in understanding further the nature of the $Si-SiO_2$ interface. It would indicate if surface reconstruction is going on, or if an oxidation reaction is going to completion. It would also help answer questions such as the value of D_{ito} or N_{f_o}; and whether it corresponds to an initial value, or is a universal constant dependent only on temperature. We are also conducting experiments to find out what effects, if any, the

235

growth conditions of the silicon dioxide layer have on the anneal time constants and the final annealed value of Q_f and D_{it}. Preliminary results indicate that the anneal kinetics have the form of an exponential decay as described above and the oxide growth conditions do have an effect on the final annealed values.

It is our intention to investigate cases of an increase in oxide charge densities after long periods of high temperature anneal. The prevalent belief is that the increse is caused by the diffusion of impurities into the quartz furnace tube, but a report which contradicts the above is that in which Si and SiO_2 are believed to react in an inert ambient forming gaseous SiO [7.7]. We may be able to learn something about the nature of the $Si-SiO_2$ interface from this part of the anneal kinetics.

The ultimate goal of our research efforts is to be able to quantitatively predict the interface charges and finally to implement the results in SUPREM. At the present time, SUPREM has no way of predicting these charges; this makes it difficult for two dimensional device simulators such as GEMINI using SUPREM outputs to be able to accurately predict threshold voltages and device transconductances. We hope that the incorporation of interface charge models will further enhance device and process simulation capabilities.

REFERENCES

[7.1]   B. E. Deal, Journal of ECS, 121, 198C, 1974.

[7.2]   D. R. Lamb and F. R. Badcock, Int. J. Electronics, 24, No. 1, 11-16.

[7.3]   R. R. Razouk and B. E. Deal, Journal of ECS, 126, No. 9, 1573.

[7.4]   W. A. Tiller, Journal of ECS, 128, No. 3, 689, 1981.

[7.5]   V. Q. Ho, T. Sugano, IEEE Elect. Dev., ED-28, No. 9, Sept. 1981, 1060.

[7.6]   F. M. Fowkes, D. W. Hess and J. Kiddon, Paper #76 at 151st Meeting ECS, May 10, 1977.

[7.7]   S. I. Raider and A. Berman, Journal of ECS, 125, No. 4, 629, Apr 1978.

[7.8]   S. P. Murarka, App. Phys. Lett., 34(9), May 1, 1979.

[7.9]   K. L. Ngai and C. R. White, J. Appl. Phys., 52(1), Jan. 1981, 320.

[7.10] J. Singh and A. Madhukar, Appl. Phys. Lett., 38(11), June, 1981.

[7.11] J. F. Grunthaner, B. F. Lewis, R. P. Vasquez and J. Maserjian, "Physics of MOS Insulators, " Raleigh, NC, June, 1980.

[7.12] J. L. Pautrat and J. D. Pfiested, Phys. Stat., Sol(a), 11, 669, 1972.

[7.13] M. Schultz, K. Blumenstock and E. Klausman, Semiconductor Silicon, 1981, pp. 463-473.

[7.14] A. Lin, D. A. Antoniadis, and R. W. Dutton, Journal of ECS, 128, No. 5, 1131, May 1981.

[7.15] A. Lin, et. al., Journal of ECS, 128, No. 5, 1121, May 1981.

[7.16] See section of this report on Point Defect Generation at the Si/SiO_2 Interface.

[7.17] T. M. Hall, Jr., "A study of the Interface Properties of Deposited Silicon Dioxide on Silicon," Engineer's Thesis, Stanford ELectronics Laboratories, Stanford University, June 1981.
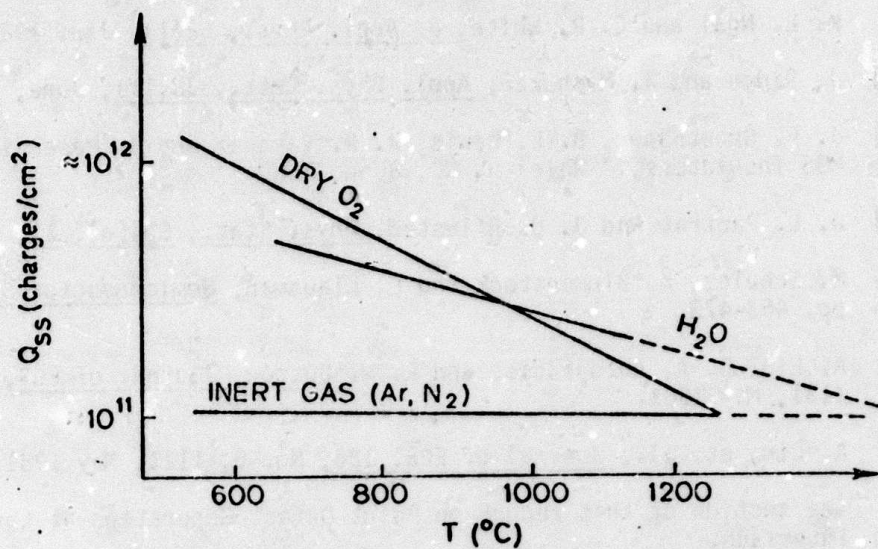
Fig. 7.1: "Q$_f$ triangle" (or Q$_{SS}$) after Deal [7.1]
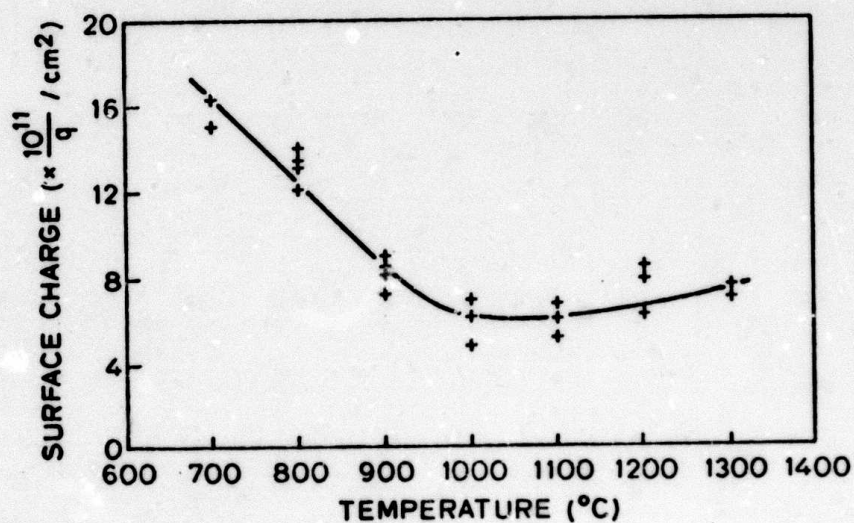
Fig. 7.2: Surface charge (Q$_f$) vs. annealing temperature in Ar after Lamb and Badcock [7.2].
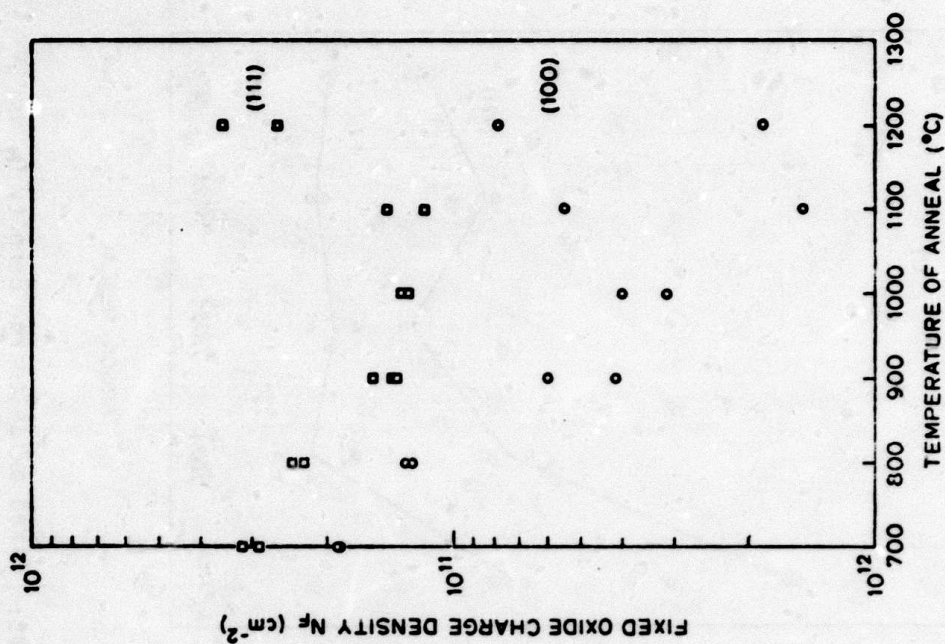
Fig. 7.4: Fixed oxide charge density Q_f vs. anneal temperature in Ar for (111) and (100) wafers.

FIXED OXIDE CHARGE DENSITY N_F (cm^{-2})

TEMPERATURE OF ANNEAL (°C)

(111)

(100)
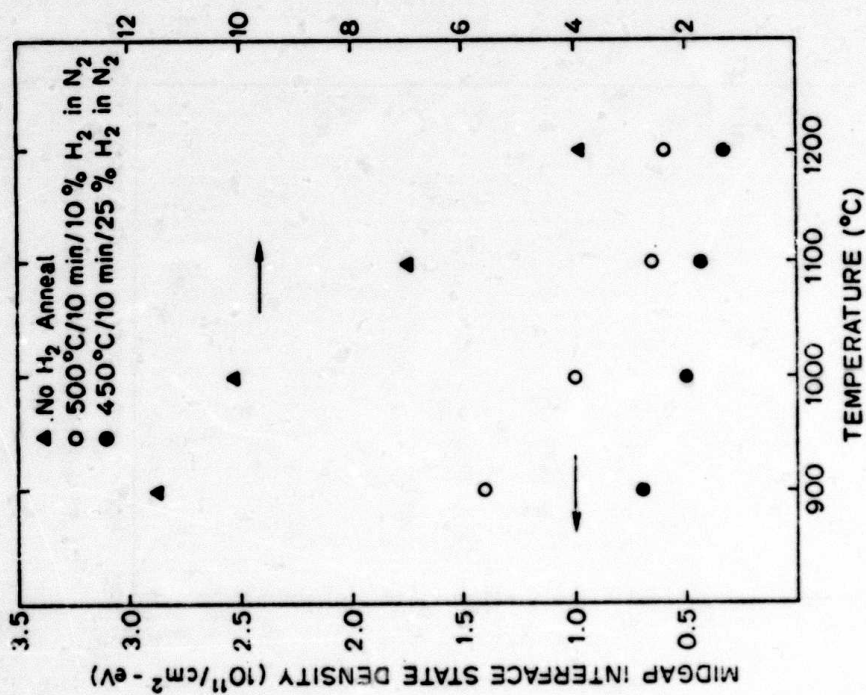
Fig. 7.3: Interface state trap density vs. growth temperature for various low temperature post metallization anneals [7.3].

MIDGAP INTERFACE STATE DENSITY (10^{11}/cm^2-eV)

TEMPERATURE (°C)

▲ No H_2 Anneal
○ 500°C/10 min/10 % H_2 in N_2
● 450°C/10 min/25 % H_2 in N_2

Fig. 7.6: Fixed oxide charge density Q_f vs. Ar anneal cycle for (111) and (100) wafers thermally cycled between 900°C and 1100°C in Ar.

Fig. 7.5: Midgap interface state density D_{it} vs. anneal temperature in Ar for (111) and (100) wafers, with no low temperature H anneal.

Fig. 7.8: As grown (fast pulled) midgap interface state density D_{it} vs. growth temperature for (111) and (100) wafers. No low temperature H anneal was used.

Fig. 7.7: As grown (fast pulled) fixed oxide charge density Q_f vs. growth temperature for (111) and (100) wafers.

8.   EFFECT OF ELECTROCHEMICAL FIELDS DURING Si OXIDATION

J. R. Rouse & C. R. Helms
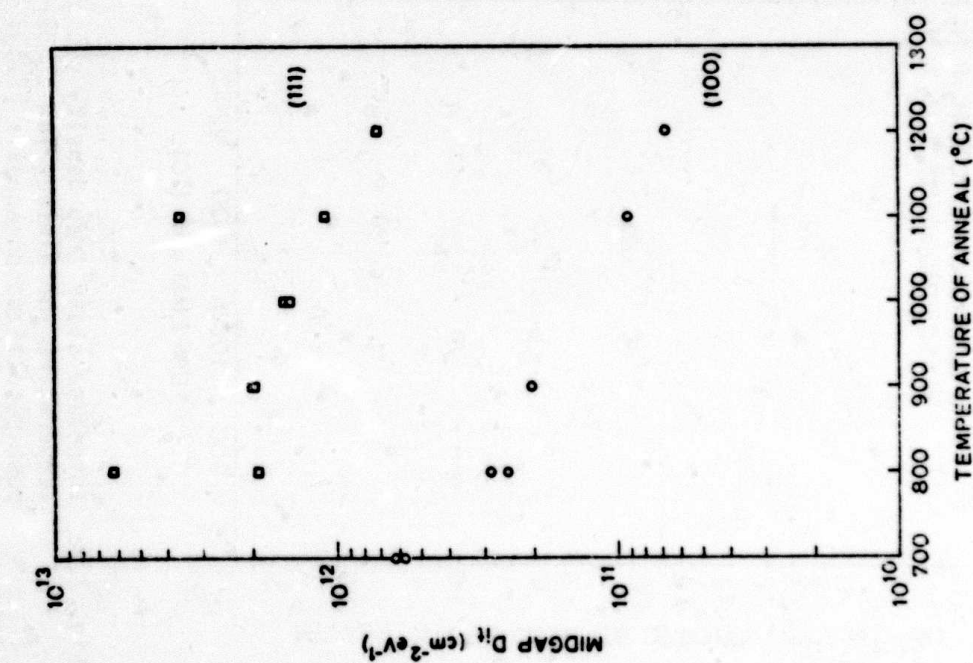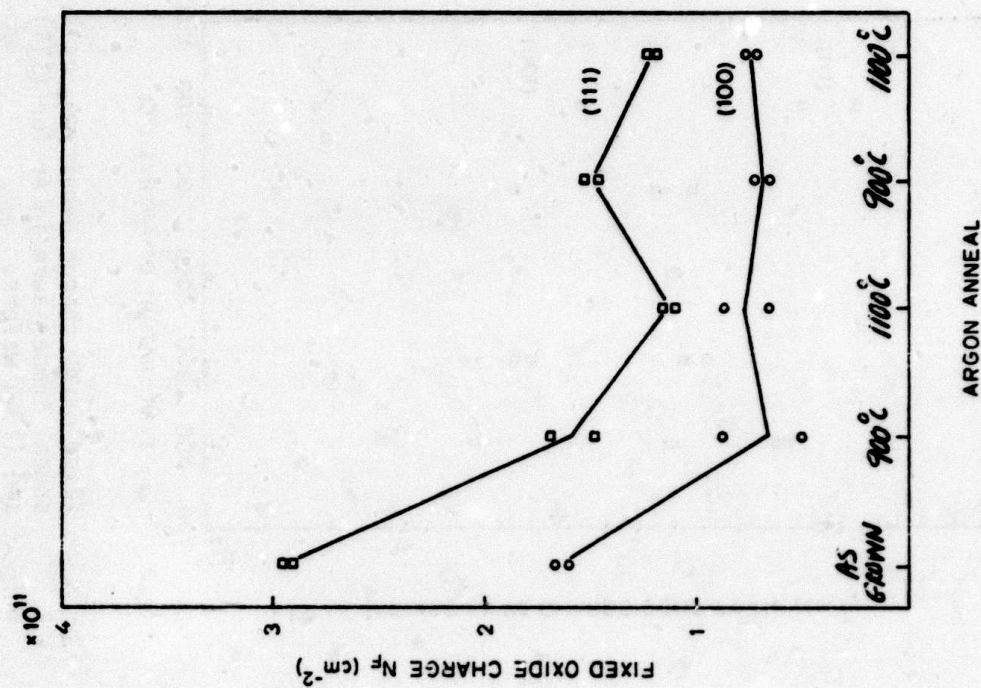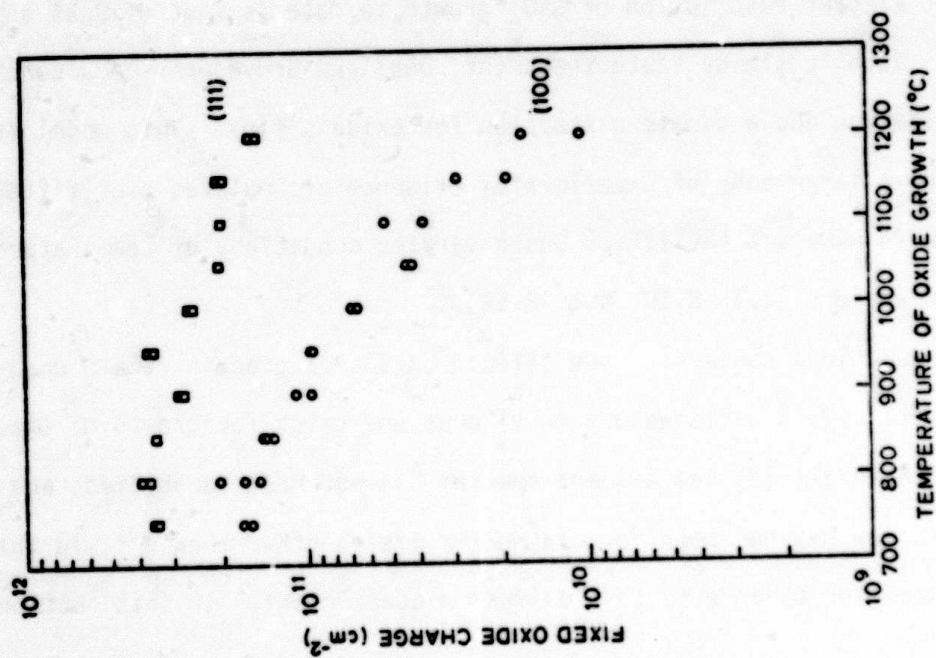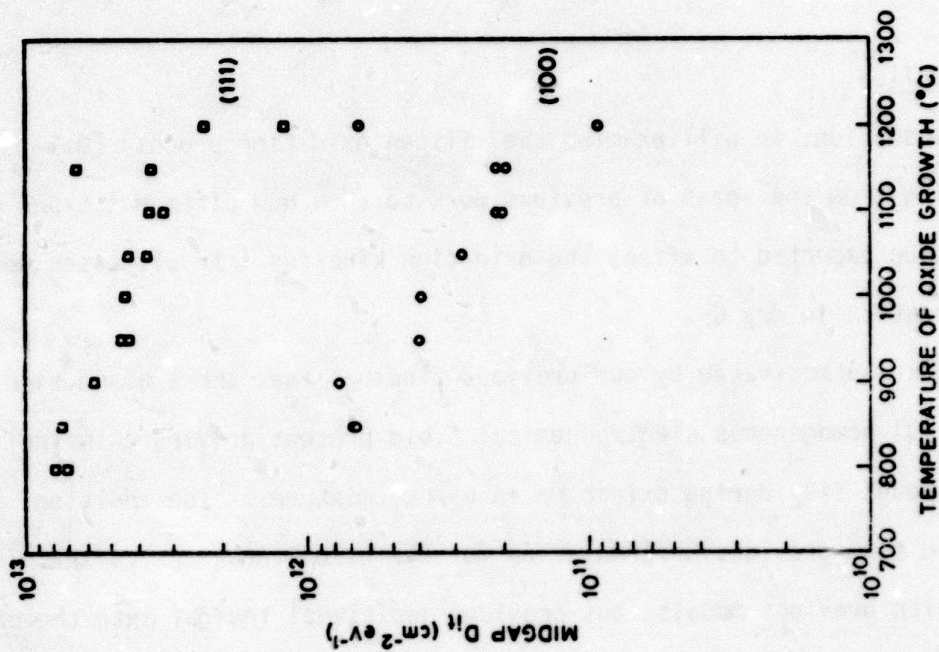
8.1   INTRODUCTION

In this section, we will examine the silicon oxidation process [8.1-8.14], drawing from the ideas of previous work to show how different types of fields might be expected to affect the oxidation kinetics. In all cases we refer to oxidation in dry O_2.

This work was motivated by our previous findings that there was a significant (0.65 eV) homogeneous electrochemical field present driving chlorine transport through SiO_2 during oxidation in O_2/HCl mixtures. The addition of this field term provides a formulation for the oxidation kinetics that is consistent with previous models, but provides additional insight into the oxidation mechanism. These ideas may be particularly important in developing models for oxidation kinetics for thin oxides necessary for future versions of SUPREM.

The most elegant description of SiO_2 growth to date is that of Deal and Grove [8.3]. Using a steady state treatment, Deal and Grove provide a concise model for oxidation and a simple expression for oxidant flux. This model is in accord with a large body of experimental evidence accumulated over fifteen years of work in numerous facilities under varying conditions of temperature and oxidation ambient [8.3, 8.10, 8.15-8.19].

Several questions concerning the silicon oxidation process remain unanswered, however: (1) a satisfactory model does not exist for growth of oxides less than 200 Å thick; (2) the oxidant species has not been identified; and (3) the growth law has not been formulated for fields other than a field that can be accounted for by an effective diffusion coefficient. In this section

242

we will address the last consideration and derive a growth law for oxidant driven by a general field.

To provide a framework for examining various oxidation models, we first present the Deal-Grove formulation of oxide growth. The Deal-Grove model [8.3] begins with the SiO_2/Si structure shown in Fig. 8.1. They assume that oxidation occurs through the flux of a single species diffusing through the oxide. Invoking steady state, they then equate flux at the boundaries and in the bulk via

$$J = h(C_{EQ} - C_o) \tag{8.1}$$

$$J = - D \frac{dC}{dx} \tag{8.2}$$

$$J = kC_i \tag{8.3}$$

C_{EQ} is the equilibrium concentration of oxidant in the oxide, and C_o and C_i are the oxidant concentrations at $x = 0$ and $x = x_o$, respectively. D is the oxidant diffusion coefficient. These equations yield

$$J = \frac{NB}{2x_o + A} = N \frac{dx_o}{dt} \tag{8.4}$$

where N is the concentration of SiO_2 molecules in the lattice, x_o is the oxide thickness, and B and B/A are the parabolic and linear rate constants

$$B = \frac{2DC_{EQ}}{N} \tag{8.5}$$

$$\frac{B}{A} = \frac{D}{k} \tag{8.6}$$

The so-called linear-parabolic growth law of Eq. (8.4) is seen to be linear for x_o small and parabolic for x_o large. The length A serves as a transition oxide thickness separating the two regimes.

243

At the oxide surface, Henry's law is applied to yield

$$C_{EQ} = Kp^q \qquad (8.7)$$

where p is the partial pressure of O_2 in the oxidizing ambient and q is some rational fraction determined by the reaction mechanisms at the surface. We can postulate numerous possibilities for the reactions which yield the oxidant species in the SiO_2 lattice. Three such reactions are

$$O_2 \rightarrow O_2 \qquad (8.8)$$

$$e^- + O_2 \rightarrow O_2^- \qquad (8.9)$$

$$4e^- + O_2 \rightarrow 2O^= \qquad (8.10)$$

Equation (8.10), for example, would yield an equilibrium surface oxidant concentration of

$$C_{EQ} \propto [O^=] = K[e^-]^2[O_2]^{1/2} \qquad (8.11)$$

These equations then would yield the respective pressure dependencies for B then as

$$O_2: \quad B = K_1 p \qquad (8.12)$$

$$O_2^-: \quad B = K_2 p \qquad (8.13)$$

$$O^=: \quad B = K_3 p^{1/2} \qquad (8.14)$$

for O_2 partial pressure p.

Deal and Grove [8.3] and Doremus [8.7] have shown that B is linearly proportional to p, and this has been born out in further studies by Deal and

244

Razouk [8.21] for 1 atm < p < 20 atm. We expect, then, that the probable oxidant species is O_2 or O_2^-.

Jorgensen [8.1] devised an experiment several years prior to the Deal-Grove work to determine the charge state of the oxidant species. The experimental arrangement is shown in Fig. 8.2. Jorgensen used porous platinum electrodes to bias the sample. He found that the side with the electrode negative with respect to the silicon experienced an accelerated oxide growth rate; and the side with the electrode positive with respect to the silicon showed retarded growth with respect to normal oxidation. He found that oxide growth became negligibly slow for a positive electrode potential of 1.62 volts. Observing that the change in free energy for the oxidation reaction is 1.76 eV at the oxidation temperature, he concluded that the applied voltage eliminated the free energy gradient for the oxidation reaction and, thus, oxidation should stop. Jorgensen concluded that these results proved that the oxidant species is negatively charged.

Raleigh [8.4] presented a compelling argument challenging Jorgensen's conclusions. He first pointed out that, in open circuit conditions, an applied electric field cannot alter the growth rate. Oxidation in open circuit conditions proceeds by the transport of an electrically neutral species. If growth proceeds by the transport of two oppositely charged species, the applied field must accelerate one particle flux and retard the other. This results in a net current flow and, hence, a charge build-up at the oxide boundaries. The resulting boundary layer of charge results in an electric field which cancels the applied field and returns the transport to the flux levels operative with no applied field. Raleigh estimated that the system would return to the no-field flux rate in the time it takes for a fractional monolayer of oxide formation.

In order for an applied electric field to affect the flux of an oxidant ion, there must therefore exist an imbalance of charge flux and, hence, a current must flow. Raleigh showed that, with his experiment, Jorgensen constructed two back to back electrochemical cells $Pt,O_2|SiO_2|Si|SiO_2|Pt,O_2$. Clearly, the presence of current flux implies that the oxidation process is altered from the normal oxidation process in open-circuit conditions. Jorgensen ignored this point in presenting his analysis. Raleigh went on to postulate the half-cell reactions:

$$O_2 + 4e^- \rightleftarrows 2O^= \qquad (8.15)$$

$$2O^= + Si \rightleftarrows SiO_2 + 4e^- \qquad (8.16)$$

In the presence of current flow, oxidation can proceed by normal diffusive flux of neutral oxygen, augmented by electrolytic SiO_2 formation on one side and decreased by electrolytic decomposition on the other. The net oxide growth is then given by

$$\frac{dx_o}{dt} = \frac{B}{2x_o} \pm \frac{etJ}{4NF} \qquad (8.17)$$

B is the parabolic rate constant, N is the concentration of SiO_2 lattice molecules, F is Faraday's constant, eJ is the total current density, and t is the ionic transport number. Using estimates of B and t, Raleigh calculated a stopping potential of 1.25 v, giving good agreement with Jorgensen's measured 1.62 v in view of the necessary approximations. In particular, the ionic transport number is not known accurately.

Implicit in Raleigh's argument is the point that the current flow provides a large population of ions which may not exist during normal oxidation.

Thus, by virtue of chemical reactions like (8.8)-(8.10), the population of charged species may be severely affected by the current flow. Thus, Jorgensen's experiment is probably not indicative of the situation which prevails during normal open-circuit oxidation. The thrust of Raleigh's paper is that experiments of the type Jorgensen performed cannot offer conclusive proof of the identity of the oxidant species. This does not preclude the possibility that the oxidant species is negatively charged. Raleigh's arguments are forceful, however, and demand that we do not insist that the oxidant species is charged.

Tiller [8.12] has given a comprehensive treatment of the general thermodynamics of the silicon oxidation process. The interface driving force for oxidation is shown to consist of the free energy necessary to generate defects and strain fields associated with oxide growth and the additional free energy needed to drive the molecular transition from oxidant particle to SiO_2. Tiller then proceeds by postulating that oxidation proceeds through the flux of $O^=$ and O_2. The oxidation kinetics are outlined in general terms:

$$\frac{dx_o}{dt} = \Omega^{SiO_2} \left[J_{O_2} + \frac{1}{2} J_{o^=} \right] \tag{8.18}$$

$$J = -\mu \frac{du}{dx} C \tag{8.19}$$

$$\frac{du}{dx} = f(\phi, \sigma, \Delta\mu_{H.M.}, \dots) \tag{8.20}$$

x_o is the oxide thickness, Ω^{SiO_2} is the molecular volume of SiO_2, μ is the mobility of a given particle species, μ is the electrochemical potential, and C is the concentration. ϕ, σ, and $\Delta\mu_{H.M.}$ denote the electric field,

247

stress field, and chemical potential due to bonding at the interface. An additional constraint is the free volume supply condition

$$\frac{dx_0}{dt} \propto J_\nu^{Si} + J^{SiO_2} \tag{8.21}$$

where J_ν^i is the vacancy flux of species i to the interface.

Tiller assumes respective gas/SiO_2 and SiO_2/Si region reactions given by

$$O_2 + 4e^- \rightleftharpoons 2O_= \tag{8.22}$$

$$v_f + Si + 2O_= + 4p^+ \rightleftharpoons SiO_2 \tag{8.23}$$

where v_f is the free volume. Using (8.22) and (8.23), Tiller concludes that $2O_=$ is equivalent to O_2 and should thus give a linear dependence of flux with O_2 pressure. Using Tiller's equations, we have

$$\frac{dx_0}{dt} = \frac{\Omega^{SiO_2}}{2} J_{O_=} = v_{O_=}[O^=] \tag{8.24}$$

By Eq. (8.22),

$$[O^=] = K[e^-]^2[O_2]^{1/2} \tag{8.25}$$

To yield the required linear pressure dependence, the species must satisfy $[e^-] = [O_2]^{1/4}$.

We have, then, at our disposal three facts of silicon oxidation with which every model must conform. First, we cannot unambiguously extract any information about the charge state of the oxidant from data in the literature at present. Secondly, whatever the precise functional form of the oxidant flux may be, it must be very closely approximated by Eq. (8.4). This is evident from the excellent agreement with the data provided by this expression.

248

Finally, the O_2 pressure dependence of the parabolic rate constant is linear. This restricts the likely candidates for the oxidant species to O_2 and O_2^-. Using these facts to guide us, we will present general arguments to extract further information about the oxidation process and will extend the Deal-Grove growth law to include oxidant transport in the presence of a constant field.

8.2 GENERAL TRANSPORT CONSIDERATIONS

As discussed in a previous section, transport through a thin film such as SiO_2 during oxidation can be described by the equation

$$J(x, t) = - D(x) \frac{\partial C}{\partial x} + \mu(x)F(x, t)C(x, t) \tag{8.26}$$

Flux J, force F, and concentration C are, in general, functions of time and position. The mobility μ and diffusion coefficient D may have some spatial dependence. The driving force F comprises numerous effects, including electric fields, stress fields, chemical bonding, defects, and activity gradients [8.12, 8.20].

In addition to the above equation, we have the continuity equation for the case of negligible recombination or generation in the oxide:

$$\frac{\partial C}{\partial t} = - \frac{\partial J}{\partial x} \tag{8.27}$$

Equations (8.26) and (8.27) give us

$$\frac{\partial C}{\partial t} = - \frac{\partial}{\partial x} \left[- D \frac{\partial C}{\partial x} + \mu FC \right] \tag{8.28}$$

We can write the condition for steady state as [8.22]

$$\frac{D}{B} \left[\frac{1}{4} \left(\frac{\phi}{k_B T} \right)^2 + \pi^2 \right] \gg 1 \tag{8.29}$$

249

where B is the parabolic rate constant of the Deal-Grove growth law. The worst case situation occurs when there is no field. In this case, Eq. (8.29) becomes

$$C_{EQ} \leq N \tag{8.30}$$

where N is the concentration of molecules in the lattice. This is clearly satisfied in any normal oxidation process, so steady state is always a good approximation.

Proceeding with the steady state approximation, we can then solve Eq. (8.26) for the oxidant distribution

$$C(x) = \frac{C(0) - J(t) \int_0^x \frac{1}{D(y)} \left[\exp - \frac{1}{k_B T} \int_0^y F(z, t) dz \right] dy}{\exp \left[- \frac{1}{k_B T} \int_0^x F(z, t) dz \right]} \tag{8.31}$$

where $J(t)$ is determined by boundary conditions.

8.3 CABRERA-MOTT FORMULATION

Before proceeding with a slightly more generalized formulation of the basic oxidation equations, it will be useful to gain a brief perspective of the historical antecedent for most present day treatments of the oxide growth process. The formulation of oxidant transport as it appears in Deal and Grove [8.3] makes use of results from Cabrera and Mott [8.13] and Bardeen, Brattain, and Schockley [8.14]. Both groups use steady state flux equations for charged species in the oxide. To illustrate this treatment, let us suppose that oxidation proceeds by the transport of O_2^- (C_i) and holes (C_h). Then, we assume an electric field is established due to the space charge produced by the differing mobilities. Cabrera and Mott make the assumption ($C_h(x) \simeq C_i(x)$), which in turn yields

$$J = -2 \frac{D_i D_h}{D_i + D_h} \frac{\partial C_i}{\partial x} \qquad (8.32)$$

This is the familiar result where the standard diffusion equation is obtained if we replace the diffusion coefficient by an effective diffusion coefficient $D_{EFF} = 2D_i D_h/(D_i + D_h)$. To obtain the electric field, we insert Eq. (8.32) into Eq. (8.26):

$$E = -\frac{k_B T}{e} \left(\frac{D_h - D_i}{D_h + D_i} \right) \frac{\partial}{\partial x} \ln C_i \qquad (8.33)$$

The solution to Eq. (8.26) then becomes

$$C(x) = C(0) - \frac{J}{D_{EFF}} x = C(0) - \left(\frac{C(0) - C(x_0)}{x_0} \right) x \qquad (8.34)$$

If we assume first order interface reaction kinetics, as is done in the Deal-Grove treatment, the flux is

$$J = kC(x_0) \qquad (8.35)$$

Inserting (8.35) in Eq. (8.34) gives

$$C(x_0) = \frac{D_{EFF}/k \; C(0)}{x_0 + D_{EFF}/k} \qquad (8.36)$$

$$J = \frac{D_{EFF} \; C(0)}{x_0 + D_{EFF}/k} \qquad (8.37)$$

which is the familiar linear-parabolic growth law.

The voltage drop across the oxide is obtained by integrating the electric field of Eq. (8.33) to obtain

$$V = \frac{k_B T}{e} \left(\frac{D_h - D_i}{D_h + D_i} \right) \ln \frac{C(0)}{C(x_0)}$$

$$= \frac{k_B T}{e} \left(\frac{D_h - D_i}{D_h + D_i} \right) \ln \left(\frac{x_0 + D_{EFF}/k}{D_{EFF}/k} \right) \qquad (8.38)$$

For thin oxides, $x_0 < D_{EFF}/k$, and the potential drop will always be much smaller than $k_B T$. Even for $x_0 \gg D_{EFF}/k$, the potential drop across the oxide varies as $\ln x_0$, so that V will still be on the order of $k_B T$. For typical growth temperatures, $k_B T \sim 0.1$ eV, corresponding to a relatively small field.

8.4 CONSTANT FIELD SOLUTION FOR OXIDE GROWTH

The general solution for a particle distribution is given in Eq. (8.31). To solve this in a specific case, we need appropriate boundary conditions for flux J. We apply the same boundary conditions used by Deal and Grove. At the surface, we have

$$J = k(C_{EQ}(0) - C(0)) \tag{8.39}$$

and

$$J = k_4 C(x_0) \tag{8.40}$$

at the interface. These are inserted in Eq. (8.31) to yield

$$J = \frac{D k_4 C_{EQ}}{\frac{D k_4}{k} + D \exp - \frac{1}{k_B T} \int_0^X F(z)dz + k_4 \int_0^X \exp - \frac{1}{k_B T} \int_0^y F(z)dz \; dy} \tag{8.41}$$

At this point, we have one fact to guide us: the Deal-Grove oxide growth relationship has been amply demonstrated to be the correct one for a wide range of pressure, temperature, oxide thickness, and cases including species such as H_2O and HCl in the oxidizing ambient [8.3, 8.10, 8.15-8.19]. This demands that oxidant flux at the SiO_2/Si interface be described by

$$J = \frac{NB}{2x_0 + A} \tag{8.42}$$

252

where N is the concentration of SiO_2 lattice molecules, and B and A are growth parameters independent of oxide thickness.

One form of the field for which Eq. (8.42) is satisfied is for

$$F(x, t) = F(t) = \frac{\phi}{x_0} \qquad (8.43)$$

where ϕ is a constant independent of x and t. In previous work we have observed such a field acting on chlorine in SiO_2 during thermal oxidation. Using this field in Eq. (8.41) gives

$$C(x, t) = C_1 + C_2 \exp (bx) \qquad (8.44)$$

$$J = \frac{P}{2x_0 + a} \qquad (8.45)$$

where

$$C_1 = \frac{J}{Db} \qquad (8.46)$$

$$C_2 = C_{EQ}(0) - \frac{J}{Db} \qquad (8.47)$$

$$b = \frac{\phi}{k_B T x_0} \qquad (8.48)$$

$$P = 2D C_{EQ}(0) \left(\frac{\phi}{k_B T}\right)\left(\frac{\beta}{\beta - 1}\right) \qquad (8.49)$$

$$\beta = \exp\left(\frac{\phi}{k_B T}\right) \qquad (8.50)$$

$$a = \frac{2D}{k_4} \left(\frac{\phi}{k_B T}\right)\left(\frac{1}{\beta - 1}\right)\left(1 + \frac{k_4}{k} \beta\right) \qquad (8.51)$$

Also, note that, for $\phi = 0$, Eq. (8.44) leads to a linear distribution.

We can rewrite Eq. (8.45) using the Deal-Grove notation, where the linear and parabolic rate constants incorporate the field. Then

253

$$J = \frac{NB}{2x_0 + A} \tag{8.52}$$

$$B = \frac{2DC_{EQ}}{N} \left(\frac{\phi}{k_B T}\right)\left(\frac{\beta}{\beta - 1}\right) \tag{8.53}$$

$$B = B_{DG} \left(\frac{\phi}{k_B T}\right)\left(\frac{\beta}{\beta - 1}\right) \tag{8.54}$$

$$\frac{B}{A} = \frac{B}{A}\bigg|_{DG} \beta \tag{8.55}$$

B_{DG} amd $(B/A)_{DG}$ are the linear and parabolic rate constants as defined by Deal and Grove. Because a homogeneous field has been shown to drive chlorine during silicon oxidation, it is attractive to conjecture that such a field also acts on the oxidant species. This possibility is made plausible by the electrochemical similarities between oxygen and chlorine. In view of this, and becuase B and B/A wculd incorporate any such field term, evaluating D, k_4, and $C_{EQ}(0)$ from oxidation data demands knowledge of ϕ.

Note that a possible field effect can have a large effect on the growth velocity. For chlorine in SiO_2 at 1000°C, we have measured $\phi = 0.65$ eV. If this field also acts on the oxidant species, then the growth rate is increased by a factor of ~1600 in the linear oxidation regime over the oxidation rate which would occur by normal diffusive flux. Thus, we see that oxide growth may be completely governed by field driven transport of oxidant particles.

Oxygen ions and holes would both be described by Eq. (8.44). This implies an exponentially increasing ion concentration and exponentially decreasing hole concentration. This, in turn, necessitates the presence of space charge in the oxide. A homogeneous field of the form in Eq. (8.43) demands zero space charge in the bulk oxide. Therefore, if charged species are present in the oxide, they must be present in low concentration. Similar con-

254

siderations have been detailed in previous work for the chlorine distribution in SiO_2.

Our general expression for field driven oxidation allows us a possible alternative explanation for Jorgensen's observed stopping potential. We rewrite the parabolic rate constant for oxidation where we assume the oxidant species is charged. We have

$$B = B_{DG}\left[\frac{\phi - \phi_A}{k_B T}\right]\left(\frac{\beta}{\beta - 1}\right) \tag{8.56}$$

$$B = \exp\left[\frac{\phi - \phi_A}{k_B T}\right] \tag{8.57}$$

where ϕ denotes the built-in potential and B_{DG} is the zero field rate constant of Deal and Grove. Jorgensen found that oxide growth was negligibly slow for an applied stopping potential of 1.62 eV at 850°C. We then have

$$\frac{B}{B_{DG}} = \left[\frac{\phi - 1.62}{0.094}\right]\left(\frac{\beta}{\beta - 1}\right) \tag{8.58}$$

In Fig. 3, we show log (B/B_{DG}) as a function of the built-in potential when the stopping potential is 1.62 eV. Jorgensen's results would be consistent with a built in potential $\phi \sim 0.7$ eV.

We note, finally, that the oxidation stopping potential need not be related to the free energy of SiO_2 formation. If our analysis of Eqs. (8.56 - 8.58) is correct, then we see that the stopping potential is determined by ϕ. This can be explained by consideration of the situation in Fig. 8.4. We schematically show the oxidant potential as a function of position in the oxide. The driving force is not determined by the free energy of formation of SiO_2. This energy is fixed and is not accessible to alteration. The potential gradient across the oxide can be varied, however. One possible means of

255

doing this is by applying an external field. If the applied field decreases the potential energy gradient, the oxidant flux will be greatly reduced. This dependence of flux on the native and applied fields is expressed through Eq. (8.56). We see, then, that Jorgensen's measured stopping voltage may provide evidence for the assertion that a potential energy gradient exists, and this drives oxidant species transport. Furthermore, the measured stopping voltage is not determined by the free energy of formation of SiO_2 but rather by the oxidant driving potential in the native oxide.

REFERENCES

[8.1] P. J. Jorgensen, "Effect of an Electric Field on Silicon Oxidation," J. Chem. Phys., 37, 874 (1962).

[8.2] J. R. Lingenza, "Silicon Oxidation in an Oxygen Plasma Excited by Microwaves," J. Appl. Phys., 36, 2703 (1965).

[8.3] B. E. Deal and A. S. Grove, "General Relationship for the Thermal Oxidation of Silicon," J. Appl. Phys., 36, 3770 (1965).

[8.4] D. O. Raleigh, "Transport Processes in the Thermal Oxidation of Silicon," J. Electrochem. Soc., 113, 782 (1966).

[8.5] F. C. Collins and T. Nakayama, "Transport Processes in the Thermal Growth of Metal and Semiconductor Oxide Films," J. Electrochem. Soc., 114, 167 (1967).

[8.6] R. Ghez and Y. J. van der Meulen, "Kinetics of Thermal Growth of Ultra-Thin Layers of SiO_2 on Silicon," J. Electrochem. Soc., 119, 1100 (1972).

[8.7] R. H. Doremus, "Oxidation of Silicon by Water and Oxygen and Diffusion in Fused Silica," J. Phys. Chem., 80, 1773 (1976).

[8.8] A. Lora-Tamayo, E. Dominguesz, E. Lora-Tamayo, and J. Llabres, "A New Model of the Thermal Growth of a Silicon Dioxide Layer," Appl. Phys., 17, 79 (1978).

[8.9] J. Blanc, "A Revised Model for the Oxidation of Si by Oxygen," Appl. Phys. Lett., 33, 424 (1978).

[8.10] C. P. Ho and J. D. Plummer, "Si/SiO$_2$ Interface Oxidation Kinetics: A Physical Model for the Influence of High Substrate Doping Levels.," J. Electrochem. Soc., 126, 1516 (1979).

[8.11] N. F. Mott, "Mechanisms for the Thermal Growth of Vitreous Oxide Layers on Silicon," Inst. Phys. Conf. Ser. No. 50, 12, (1980).

[8.12] W. A. Tiller, "On the Kinetics of the Thermal Oxidation of Silicon," J. Electrochem. Soc., 127, 619 (1980).

[8.13] N. Cabrera and N. F. Mott, "Theory of the Oxidation of Metals," Rep. Prog. Phys., 12, 163 (1948).

[8.14] J. Bardeen, W. H. Brattain, and W. Schockley, "Investigation of Oxidation of Copper by Use of Radioactive Cu Tracer," J. Chem. Phys., 14, 714 (1946).

[8.15] K. Hirabayashi and J. Iwamura, "Kinetics of Thermal Growth of HCl-O$_2$ Oxides on Silicon," J. Electrochem. Soc., 120, 1595 (1973).

[8.16] Y. J. van der Meulen and J. G. Cahill, "Effects of HCl and CL$_2$ Additions on Silicon Oxidation Kinetics," J. Electronic Mat., 3, 371 (1974).

[8.17] D. W. Hess and B. E. Deal, "Kinetics of the Thermal Oxidation of Silicon in O$_2$/HCl Mixtures," J. Electrochem. Soc., 124, 735 (1977).

[8.18] C. P. Ho, J. D. Plummer, J. D. Meindl, and B. E. Deal, "Thermal Oxidation of Heavily Phosphorus Doped Silicon," J. Electrochem. Soc., 125, 665 (1978).

[8.19] B. E. Deal, D. W. Hess, J. D. Plummer, and C. P. Ho, "Kinetics of the Thermal Oxidation of Silicon at 1200°C," J. Electrochem. Soc., 125, 339. (1978).

[8.20] A. T. Fromhold, Jr., Theory of Metal Oxidation, Vol. I, North-Holland (1976).

[8.21] B. E. Deal and R. R. Razouk, Stanford Integrated Circuits Laboratories Report, 118 (1981), to be published.

[8.22] J. W. Rouse and C. R. Helms, to be published.

[8.23] B. E. Deal, "The Current Understanding of Charges in the Thermally Oxidized Silicon Structure," J. Electrochem. Soc., 121, 193C (1974).
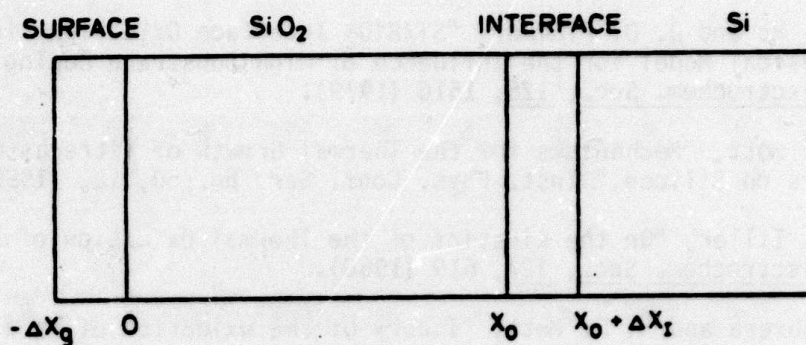
SURFACE   SiO$_2$ INTERFACE   Si

$-\Delta x_g$ O x_0 $x_0 + \Delta x_I$

Fig. 8.1: Schematic of the SiO$_2$/Si structure. Surface and interface regions differ in structure from the bulk oxide.
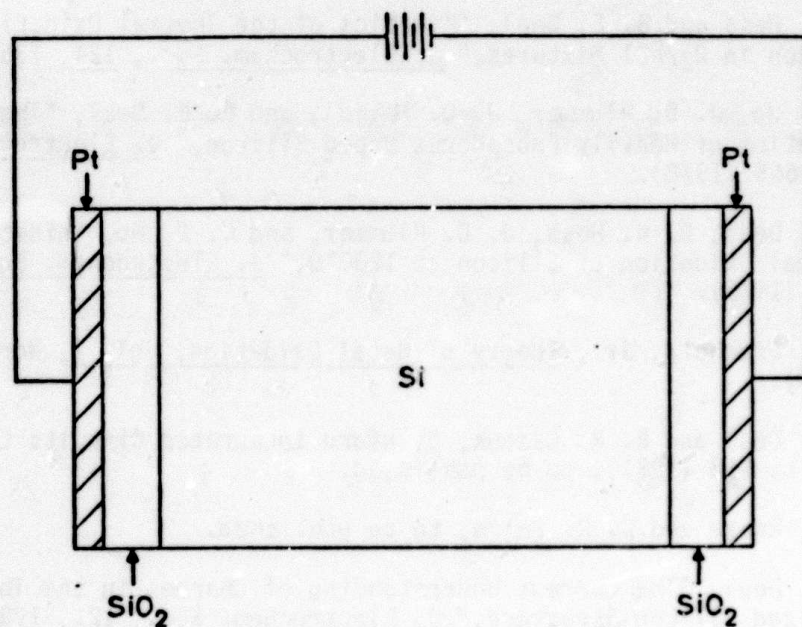

Pt   Pt

Si

SiO$_2$ SiO$_2$

Fig. 8.2: Jorgensen's experimental arrangement for studying the effect of an electric field on silicon oxidation.
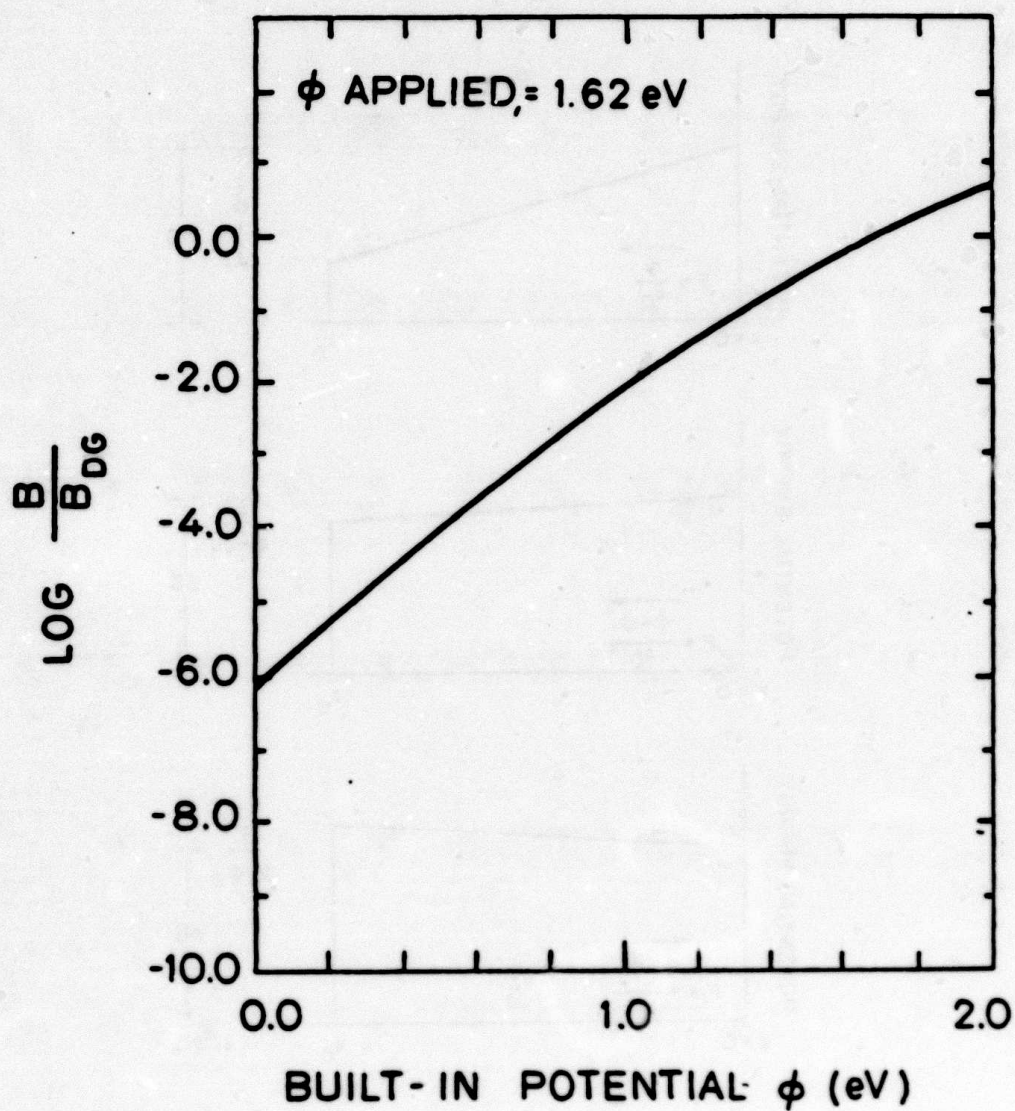
Fig. 8.3: Normalized parabolic rate constant as a function of
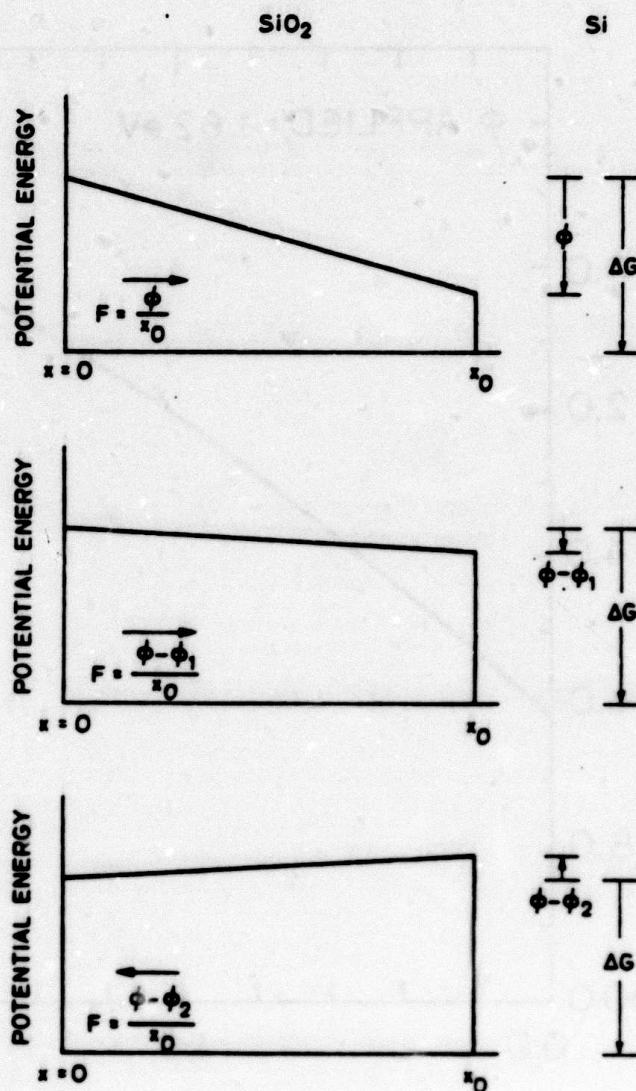built-in oxide potential for an applied retarding
potential of 1.62 eV.

Fig. 8.4: Schematic diagram of the oxidant potential energy as a function of position in the oxide. The oxidant flux in the bulk is determined by this field rather than the free energy of formation of SiO_2 (ΔG).

9. EXCHANGE OF THE OXIDANT WITH SiO$_2$ DURING THERMAL OXIDATION IN DRY O$_2$

J. W. Rouse, C. R. Helms, C. J. Han

9.1 INTRODUCTION

In this section we report on measurements of the exchange of the diffusing oxidant with SiO$_2$ lattice during the thermal growth of SiO$_2$ in O . These experiments were performed on samples similar to those reported on by Rosencher et. al. [9.1] but using more sensitive techniques so that levels of exchanged lattice oxygen could be detected. We will use the term lattice oxygen throughout this chapter for oxygen atoms which are part of the SiO$_2$ network and bonded to two Si atoms. This is different from the diffusant species for dry oxidations which is most certainly interstitial O$_2$ molecules.

9.2 EXPERIMENTAL RESULTS AND DISCUSSION

To achieve better sensitivity we grew thick oxides ($\approx$.2000 A) in enriched ^{16}O$_2$ to reduce the background ^{18}O$_2$ content by about a factor of 10 over its natural abundance. Thin ($\approx$100 A) additional oxides were then grown in ^{18}O$_2$. The ^{18}O profiles were then measured using SIMS. Using this combination of low ^{18}O background in the initial oxides and the superior sensitivity of SIMS we hoped to detect the rate of exchange of the diffusion ^{18}O$_2$ with the previously grown Si^{16}O$_2$.

In previous experiments of this type [9.1] little or no exchange of the diffusing O$_2$ with the previously formed lattice was observed, except at the surface of the previously grown films. This appears to be inconsistent with exchange results for vitreous silica [9.2-9.5] but, as recently pointed out by Revesz and Schaeffer [9.6], the conditions of the early exchange experiments during oxidation produced a system where very little exchange may have been expected.

Our results are consistent with a diffusion limited SiO_2 growth process for thick oxides were the dominant mechanism is the diffusion of molecular O_2 (nothing being said about its charge state) through the SiO_2. This molecular oxygen is free to exchange with the lattice but the rate of exchange is so slow for the growth times normally employed in silicon oxidation as to be insignificant when considering the oxidation kinetics. In addition to this transport process lattice diffusion is also occurring, but at such a slow rate as to be negligible. To investigte the diffusion process we have chosen conditions for Si oxidation in dry O_2 which would be similar to a permeation experiment [9.7]. To

To approximate these conditions we grew the initial $Si^{16}O_2$ to a thickness at which the growth rate approached the parabolic limit.

In our experiments the initial oxides were grown to a thickness of 2200 A at 1150°C. At this temperature the ratio of the parabolic to linear rate (given by A/2 in the Deal-Grove format) [9.8] ranges from a factor of 3 to 7 depending on the value of A chosen.* This is well into the parabolic regime for the whole possible range. So as to have the best chance of measuring the exchanged ^{18}O the initial $^{16}O_2$ oxidation was performed in purified oxygen which contained 0.02 at $^{18}O_2$ [9.11] which is enriched by a factor of ten over its natural abundance. The oxygen profiles were measured by SIMS using Cs^+ ions at 948 KeV [9.12]. A profile of a control $^{16}O_2$ oxide is shown in Fig. 9.1 where both the ^{16}O and ^{18}O backgrownd levels are indicated. The background level of ^{18}O is 0.05% of the

---

* Values for A are determined by a fit to the relationship $X^2 + Ax = B(t + \tau)$ where A, B and τ.are adjustable parameters. The choice of A is very sensitive to the choice of τ; previous work [9.9] would give $A = 1.05 \times 10^5$ for our conditions ($\tau = 204$ sec). New analysis [9.10] however, gives better fits for $\tau = 20.7$ sec. and $A = 4.5 \times 10^{16}$ cm, thus the range of 3 to 7 quoted above.

^{16}O concentration which is higher than the abundance in the ^{16}O used. This is possibly due to exchange with the quartz oxidation tube used for these experiments, or impurities in the SIMS system during measurement. The sequential oxidations were performed as follows: the furnace tube was outfitted with an end cap with the push rod extending from a small opening at its end. The samples (one control, one for the sequential oxidation) were inserted in to the flowing $^{16}O_2$ at the oxidation temperature. After 80 minutes the samples were pulled and cooled in flowing $^{16}O_2$ for 15 minutes. The control was then removed and the system purged in N_2 for 5 min. The sequential sample was then inserted. After 5 minutes the flow was switched to $^{18}O_2$ for 8 minutes for the sequential oxidation step. Using values of $B = 9.48 \times 10^{14}$ cm^2/sec, $A = 4.51 \times 10^{-6}$ cm, and $\tau = 10.7$ sec [9.10] yields a value of thickness for the 80 minutes of oxidation of 1924 Å where B is the parabolic rate constant, B/A the linear rate constant, and τ the initial transient time. This compares to a value of 2200 Å measured ellipsometrically. The additional 8 minutes of oxidation in $^{18}O_2$ was not all at atmospheric pressure due to the sequence of steps followed. Using a plug flow model would give an equivalent of 6 minutes of oxidation at atmospheric pressure (perfect mixing would give a 4.2 minute atmospheric pressure time product). The additional oxide grown should follow the relationship

$$\Delta X = \frac{Bt}{2X_o + A} = 60 \text{ Å} \tag{9.1}$$

for a 5 minute oxidation time (assuming no lag time through the oxide).

A SIMS profile for this sample is shown in Fig. 9.2. In addition to the $Si^{18}O_2$ grown at the interface and the large surface concentration of $Si^{18}O_2$ previously reported [9.1], significant exchange with the $Si^{16}O_2$ lattice is observed. This is seen as the ^{18}O concentration profile extending through the ^{16}O oxide as well as an apparent shoulder at the $Si^{16}O_2/Si^{18}O_2$ interface. To verify that indeed the ^{18}O observed was exchanged and part of the lattice (not

trapped interstitial oxygen, for example) samples were further subjected to both 10 min N_2 anneal and 10 min $^{16}O_2$ oxidations at the 1150°C growth temperatures used. No differences in ^{18}O profiles through the ^{16}O lattice were observed for either of these treatments.

A few comments concerning the SIMS profile are in order. First, is the question of depth resolution. The FWHM of the ≈60 Å of additional oxide grown for the sample of Fig. 9.2 is 210 Å. This is due to the inherent broadening factors, associated with the ion damage profile plus other artifacts in the SIMS experiments. The slope of the exchanged profile, through the $Si^{16}O_2$ however, is small enough so as to be little perturbed by this effect. The peak in the ^{18}O profile in the surface will also show this effect but possibly to a lesser extent.

The shapes of the profiles are also affected by other SIMS artifacts which lead to some variation in sensitivity factor with depth. For example, notice the dip in both ^{16}O and ^{18}O profiles at a depth of 1500 Å. This effect is reproducible, but probably due to the proximity of the Si/SiO_2 interface. Therefore, fine features in the profiles, especially near interfaces, should not be taken too seriously.

With all this in mind we can, however, deduce a considerable amount about the transport of the oxidizing species from the data shown. We start by considering two species. Of interest are O_2 that diffused rapidly through the SiO_2 network in an interstitial form, and O that is part of the lattice and diffuses at a much slower rate [9.6]. For example, the permeation data of oxygen through vitreous silica would give a value of D ≈10^{-8} cm^2/sec for the fast diffusing O_2 molecules [9.7] whereas data on oxygen exchange with vitreous silica would give a value of D = ≈10^{-14} cm^2/sec for slower diffusing lattice oxygen [9.3]. For the conditions of our experiment, the additional $Si^{18}O_2$

that is grown is due entirely to the first mechanism, the $^{18}O_2$ diffusing through the $Si^{16}O_2$ without significant interaction with the lattice. This can easily be seen if we compare the lag time for these two processes for the 2200 A oxide thickness used. We obtain $\tau = 5 \times 10^{-3}$ and 5×10^3 sec for the molecular and lattice transport respectively.

The expected profile of the interstitial diffusing molecular $^{18}O_2$ associated with the molecular diffusion mechanism given by $C_o \left(1 - \dfrac{x}{\dfrac{A}{2} + xo} \right)$ is shown in Fig. 9.3 as the solid line [9.1].

If we assume exchange from the diffusing O_2 takes place via the reaction

$$^{18}O_2)_I + {}^{16}O_L \rightleftharpoons {}^{16}O^{18}O)_I + {}^{18}O_L \qquad (9.2)$$
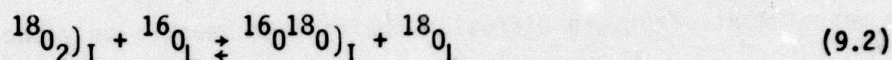
where the I and L subscripts refer to interstitial and lattice forms respectively then the ^{18}O concentration can be given by

$$C^{18} \simeq C_I{}^{18} = k \int C_I{}^{18} dt$$

$$\simeq C_I{}^{18} kt \qquad (9.3)$$

where k is the first order rate constant for the reaction in the forward direction and kt >> 1 for our conditions.

Our data is shown in Fig. 9.3 as the dashed line as well a fit to this relationship with k = 5.01 sec^{-1} as the dashed dotted line (using 5 minute time for the sequential oxidation step).

As mentioned above, in addition to this exchange mechanism there will be lattice diffusion from the surface as well as back from the $Si^{18}O_2$ grown at the interface. The peak in surface ^{18}O is due either to a higher exchange rate at the surface than in the bulk of the oxide or a higher surface concentration of the interstitial form. This being the case, the lattice ^{18}O that has exchanged at the surface will diffuse into the ^{16}O oxide. Similar mixing will occur at

265

the $Si^{16}O_2/Si^{18}O_2$ interface. This may be responsible for the shape of the ^{18}O profile near the surface as well as the shoulder in the ^{18}O profile at the $Si^{16}O_2/Si^{18}O_2$ interface due to a slow diffusion front of lattice oxygen. However, the SIMS artifacts mentioned above make this interpretation unclear.

Our results reported here show definitely the relationship between the previous determinations of diffusion coefficient in vitreous silica. The permeation experiments performed by Norton [9.7] are closely related to silicon oxidation in that the dominant mechanism in both O_2 permeation and Si Oxidation is the diffusion of molecular O_2 through the SiO_2 lattice with little exchange. The exchange measurement of oxygen diffusion coefficient measure only the lattice component which is a very small fraction of the total flux.

One difficulty with these measurements is the possible effect of small concentration of water vapor that may be present. Recent studies [9.13, 9.14] have shown that wafer vapor can increase the exchange rate by many orders of magnitude over what we observe here. We estimate, however, that 100 ppm water vapor concentrations would be required to generate the exchange rate we observe, whereas the water vapor concentration in our experiment was approximately 10 ppm or less. In addition, the subsequent N_2 anneal and $^{16}O_2$ oxidation performed on the samples would cause additional changes in the ^{18}O profiles if water vapor were a problem. Since these treatments, however, produced no additional movement of the exchanged ^{18}O we feel confident that no water vapor artifacts were important in these experiments.

REFERENCES

[9.1]   E. Rosencher, A. Straboni, S. Rigo, G. Amsel, Pll. Phys. Lett. 34, 255 (1979).

[9.2]   E. W. Sucov, J. Am. Ceram. Soc., 46, 14 (1963).

[9.3]   R. Haul and G. Dumbgen, Z. Electrochem. 66, 636 (1962).

[9.4]   E. L. Williams, J. Am. Ceram. Soc., 48, 190 (1965).

[9.5]   K. Muehlenbachs and H. A. Schaeffer, Am. Mineral, 15, 179 (1977).

[9.6]   A. G. Revesz, H. A. Schaeffer, J. Electrochem. Soc., 129, 357 (1982).

[9.7]   R. J. Norton, Nature, 171, 701 (1961).

[9.8]   B. E. Deal and A. S. Grove, J. Appl. Phys., 48, 2891 (1977).

[9.9]   B. E. Deal, D. W. Hess, J. D. Plummer, C. P. Ho, J. Electrochem. Soc., 125, 339 (1978).

[9.10] L. N. Lie, R. R. Razouk and B. E. Deal, to be published.

[9.11] Obtained from Monsanto Mound Hill facility

[9.12] SIMS Analysis performed at C. A. Evans and Associates.

[9.13] R. Pfeffer and M. Ohring, J. Appl. Phys., 52, 777 (1981).

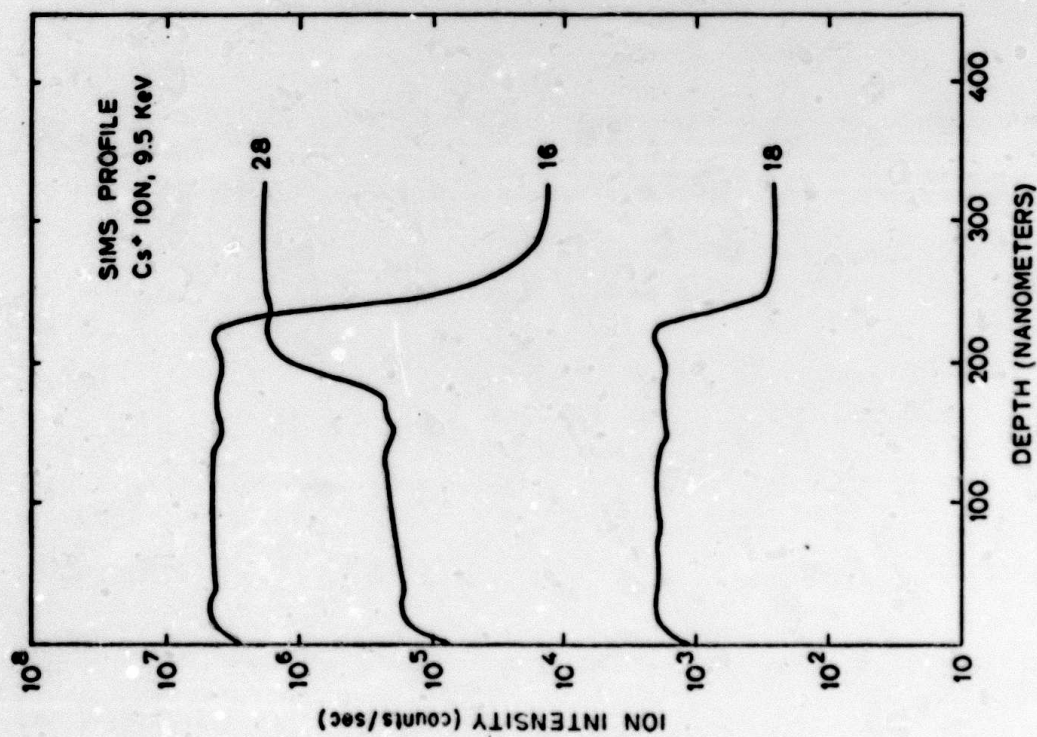[9.14] J. C. Mikkelsen, J. Appl. Phys. Lett., 40, 336 (1982).

Fig. 9.2: Sims profile of the sample of figure 9.1 with an additional ≈100 Å of $Si^{18}O_2$ grown.

Fig. 9.1: Sims profile of 2200 Å of SiO_2 grown on Si at 1150°C in $^{16}O_2$. Mass 18 signal is the ^{18}O background.
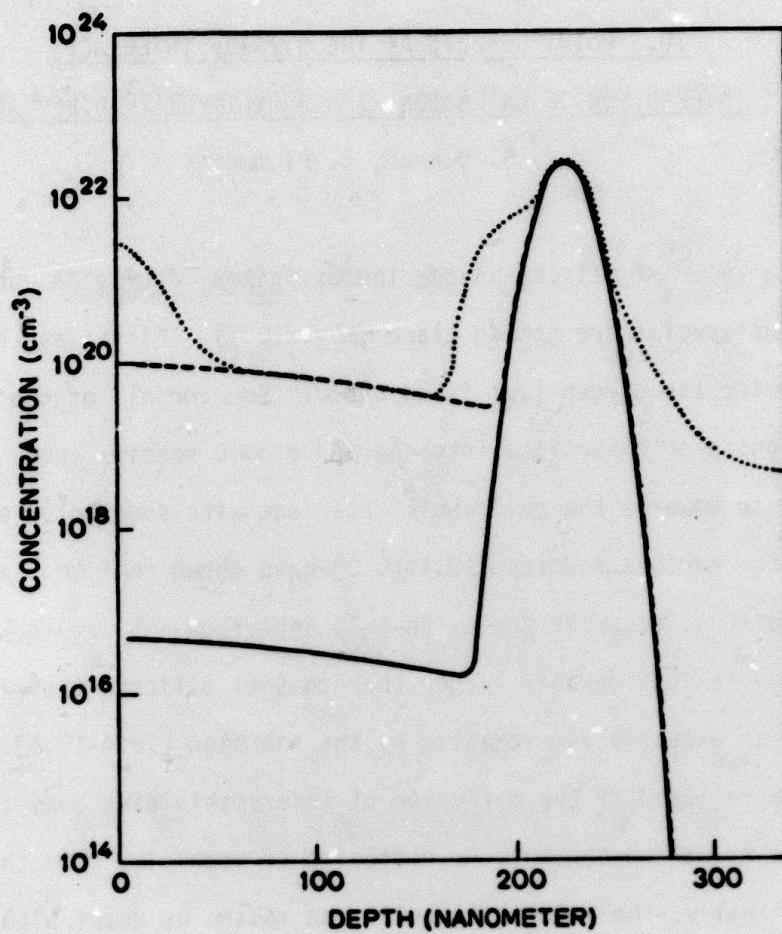
Fig. 9.3: ^{18}O profile of figure 9.2 (dotted line)
along with expected profile if no exchange
were occuring (solid line) and fitted
profile with exchange included (dashed line).

269

10. POINT DEFECTS AT THE Si/SiO$_2$ INTERFACE;

INTERACTION OF OXIDATION WITH BULK IMPURITY DIFFUSION

S. Dunham, J. Plummer

Within a wafer of silicon undergoing oxidation, diffusion and reaction of many different species are taking place (Fig. 10.1). First, at the oxide-gas interface, molecular oxygen (O$_2$) is adsorbed. Some or all of the molecular oxygen may ionize or dissociate into charged atomic species (O^X). All of these species diffuse towards the oxide-bulk interface with some total effective diffusion constant. Various studies [10.1-10.3] have shown that the oxidizing reaction, or reactions, occur at the oxide-bulk interface. Since the molecular volume of SiO$_2$ is considerably larger than that of silicon, there is a lattice mismatch and an excess volume created by the reaction [10.4-10.6]. The resultant stress can be relieved by the diffusion of interstitialcies away from the interface into the oxide and the bulk or diffusion of vacancies from the bulk to the interface. Finally, the interstitialcies and vacancies react with each other in the bulk, creating and destroying Frenkel pairs, while the silicon interstitialcies that flow into the oxide react with the incoming oxygen species to complete the oxidation process.

It is these point defect fluxes and reactions that are the major sources of interaction between oxidation rates and the properties of the oxide and bulk silicon. As the rate of oxidation increases, the fluxes of point defects must also increase to relieve the stress. The densities of point defects near the interface move farther from equilibrium at the same time, resulting in an increase in the numbers of interstitialcies and a decrease in the numbers of vacancies. Since impurities diffuse by way of point defect mechanisms, impurity diffusion constants are either enhanced (OED) [10.7-10.11] or reduced (ORD) [10.12] during

oxidation, depending on whether the impurities diffuse preferentially with inter-stitialcies or vacancies, respectively. Therefore, oxidation reshapes expected doping profiles. Concurrently, stacking faults, which grow by absorbing inter-stitialcies and shrink by absorbing vacancies, tend to grow during oxidation rather than shrink, as they do in inert ambients (OISF) [10.9,10.13,10.14-10.17].

The impurity density in silicon determines the Fermi level, which in turn influences the equilibrium point defect concentrations. As Ho and Plummer observed [10.18,10.19], an increase in the equilibrium vacancy concentration pro-duces an increased flux of vacancies toward the interface, speeding up the oxida-tion process. Thus we can see that point defect fluxes both affect and are affected by the rate of oxidation.

The research objective then is to tie all these associated phenomena togeth-er in a unified model. At the interface there are two probable reactions occurring as silicon is oxidized.

$$O_I + Si - Si \rightarrow Si-O-Si \qquad (10.1)$$

and

$$O_I + V_{Si} \rightarrow O_S \qquad (10.2)$$

O_I represents a general interstitial oxygen species, molecular or atomic, and O_S an oxygen atom in a substitutional site. We have ignored for the moment the presence of an atomic oxygen species as a product of the above reactions, with molecular oxygen species as reactants. Note that both reactions are non-stoich-iometric in SiO_2, they are merely meant to illustrate the oxidation mechanism.

Digressing for the moment, we observe that according to physical and theor-etical calculations [10.20], under most conditions the equilibrium form of a "silicon interstitial" is actually what should be called (as we have done) an interstitialcy. In other words, interstitial silicon atoms do not sit unbonded in the crystal structure "gaps" (Fig. 10.2), but rather, depending on their

271

ionization state and other conditions, either share a site with another atom
(Fig. 10.3) or lie in a bond-centered site (Fig. 10.4). Among the implications
of this is a large decrease in the expected activation energy of Frenkel (vacan-
cy-interstialcy) pairs, since less than 4 Si-Si bonds have to be broken. There-
fore, both vacancies and interstitialcies can be present in significant numbers
at equilibrium.

Going back to our oxidation reactions, reaction (10.1) consists of the
breaking of a Si-Si bond and the insertion of an oxygen atom between the two
silicon atoms (Fig. 10.6). This reaction should occur at a rate proportional to
the oxygen density at the interface ($[O]_o$), the number of surface sites
(N_{surf}) and the exponential of the activation energy required to break the
Si-Si bond and insert an oxygen atom (E_1):

$$\frac{dx}{dt}\Big|_1 \propto [O]_o \, N_{surf} \, \exp(-E_1/kT) \tag{10.3}$$

Thinking in terms of interstitialcies, we observe that this is similar to the
structure of a bond-centered interstitialcy, and could be the first step in the
creation of a silicon interstiticialcy (as in $O_I + Si \rightarrow O_s + Si_I$). This
interstitialcy can diffuse into the bulk (Fig. 10.5) or into the oxide and thus
relieve the stress caused by inserting oxygen atoms.

Reaction (10.2) involves the insertion of an oxygen atom into a vacant sili-
con site (Fig. 10.6). The rate of oxidation by the vacancy mechanism should be
proportional to the density of oxygen species at the interface, the percentage of
silicon sites that are vacant, the total number of interface sites, and an effec-
tive diffusion constant (D_{eff}) that determines how frequently the vacancies and
oxygen species come together.

$$\frac{dx}{dt}\Big|_2 \propto [O]_o \, \frac{[V]_o}{[Si]} \, N_{surf} \, D_{eff} \tag{10.4}$$

272

$[V]_o$, the interface vacancy concentration, depends on bulk diffusion, Fermi level and Frenkel pair creation rate. The total oxidation rate should be the sum

$$\frac{dx}{dt} = \frac{dx}{dt}\bigg|_1 + \frac{dx}{dt}\bigg|_2 \qquad (10.5)$$

The resulting structure after reaction (10.2) is the same as with reaction (10.1) after the interstitialcy has diffused away (Figs. 10.5 and 10.6). The Si-O-Si bonds are larger than the Si-Si bonds but smaller than two Si-Si bonds so a combination of reactions with and without point defect interactions could result in an approximate lattice fit (Fig. 10.7). In order to provide such a fit, a certain percentage of oxidation reactions should result in point defect generation/annhilation. Therefore, the ratio between point defect fluxes at the interface and the rate of oxidation should be a constant depending on the silicon orientation.

In lightly-doped silicon, the equilibrium density of vacancies is quite small (10^{14} cm^{-3} at 1100°C) so that the vacancy mechanism has negligible effect. This is affirmed by the evidence that high doping ($>10^{19}$ cm^{-3} at 900°C) is required before any enhancement of the oxidation rate occurs [10.18,10.19].

Therefore $dx/dt = dx/dt\big|_1$ and the total point defect flux is approximately the interstitialcy flux (I_{flux}). Combining this with (10.3):

$$I_{flux} \propto \frac{dx}{dt} \propto [O]_o \qquad (10.6)$$

It has been observed that in lightly-doped silicon, oxidation is more than 99% complete [10.21], so virtually all of the silicon interstitialcies created by oxidation must diffuse into the oxide. The interstitialcies react with the incoming oxygen as they diffuse away from the oxidizing interface. The continuity equation for interstitialcies in the oxide is:

$$\frac{d[I]}{dt} = D_I^{SiO_2} \frac{d^2[I]}{dx^2} - k_I [I][0] \qquad (10.7)$$

Near steady-state $dI/dt \simeq 0$. Also assume for the moment that $[0] \simeq [0]_o$,, a constant, near the interface. (This means the analysis will be most valid in the linear regime of oxidation.)

$$\frac{d^2[I]}{dx^2} = \frac{k_I}{D_I^{SiO_2}} [0]_0 [I] = k' [0]_0 [I] \qquad (10.8)$$

Equation (10.8) has solutions of the form:

$$[I] = A \exp[(k' [0]_0)^{1/2} y] + B \exp [-(k'[0]_0)^{1/2} x] \qquad (10.9)$$

The boundary condition at the interface is that $[I]_{x=0} = X[I]_0$ where X is the segregation coefficient for interstitialcies and $[I]_0$ is the density of interstitialcies in the bulk. We know that there is little silicon transfer through the oxide since the reaction takes place almost totally at the oxide-bulk interface [10.1-10.3]. Therefore $A \simeq 0$ and $B \simeq X[I]_0$. By Fick's law, the flux at $X=0$ is:

$$I_{flux} = -D_I^{SiO_2} \frac{d[I]}{dx} \Big|_{x=0} = D_I^{SiO_2} (k'[0]_0)^{1/2} X[I]_0 \qquad (10.10)$$

Combining with (10.6)

$$[I]_0 \propto [0]_0^{1/2} \propto \frac{dx}{dt}^{1/2} \qquad (10.11)$$

This approximately agrees with OED and OISF experimental results which show a $(dx/dt)^n$ dependence, where $m \simeq 0.5$ [10.9-10.12, 10.15-10.19] since OED and OISF should be proportional to [I] in the bulk. According to Hill [10.22] OED and thus [I] is nearly constant for 20 μm into the silicon. Therefore for most regions of interest, $[I] \simeq [I]0$. Modifying our assumption that $[0] = [0]_o$

274

and doing numerical analysis will change the exact power relationship in (10.11). In particular, an analysis which takes into account the parabolic regime of oxidation should be expected to give a different power law relationhip.

Physically, silicon interstitialcies plus SiO_2 is the same as SiO_x with $X < 2$. This model therefore fits with observations of nonstoichiometry near the interface. Specifically, the parallels are easy to see between this model and Tiller's [10.4-10.6] interface layer of SiO_x. The predictions he makes should hold qualitatively for a nonuniform layer such as in this model.

The aim of further work is to more accurately define the model and to determine the reaction rates and constants so that quantitative predictions can be made. There are two steps to these results. First, in-depth calculations must be done, including numerical integration of the nonlinear continuity equations that result from applying this model. Second, experiments will test the predictions and define the fundamental constants needed for quantitative results. One such experiment is underway to determine whether the vacancies supposedly used up in oxidation of highly-doped material are created thermally throughout the bulk silicon or just near the oxidizing interface.

REFERENCES

[10.1] M. M. Atalla, in Properties of Elemental and Compound Semiconductors, 1960, ed. by H. Gatos, Interscience, 163 (1960).

[10.2] J. R. Ligenza and W. G. Spitzer, J. Phys. Chem. Solids, 14, 131 (1960).

[10.3] P. J. Jorgensen, J. Chem. Phys., 37, 874 (1962).

[10.4] W. A. Tiller, J. Electrochem. Soc., 127, 619 (1980).

[10.5] W. A. Tiller, J. Electrochem. Soc., 127, 625 (1980).

[10.6] W. A. Tiller, J. Electrochem. Soc., 128, 689 (1981).

[10.7] D. A. Antoniadis, A. G. Gonzalez, and R. W. Dutton, J. Electrochem. Soc., 27, 813 (1978).

[10.8] R. Francis and P. S. Dobson, J. Appl. Phys., 50(1), 280 (1979).

[10.9] S. M. Hu, J. Appl. Phys., 45, 1567 (1974).

[10.10] K. Taniguchi, K. Kurosawa, and M. Kashiwagi, J. Electrochem. Soc., 127, 2243 (1980).

[10.11] A. M. Lin, D. A. Antoniadis, and R. W. Dutton, J. Electrochem. Soc., 128, 1131 (1981).

[10.12] S. Mizuo and H. Higuchi, Jap. J. Appl. Phys., 20, 739 (1981).

[10.13] R. B. Fair, J. Electrochem. Soc., 128, 1360 (1981).

[10.14] S. P. Murarka, J. Appl. Phys., 48, 5020 (1977).

[10.15] A. M. Lin, R. W. Dutton, D. A. Antoniadis, and W. A. Tiller, J. Electrochem. Soc., 128, 1121 (1981).

[10.16] S. P. Murarka, Physical Review B, 16, 1849 (1977).

[10.17] S. M. Hu, J. Appl. Phys., 51, 3666 (1980).

[10.18] C. P. Ho and J. D. Plummer, J. Electrochem. Soc., 126, 1516 (1979).

[10.19] C. P. Ho and J. D. Plummer, J. Electrochem. Soc., 126, 1523 (1979).

[10.20] J. W. Corbett and J. C. Bourgoin, in Point Defects in Solids, Vol. 2, 1975, ed. by J. H. Crawford, Jr. and L. M. Slifkin, Plenum, 11 (1975).

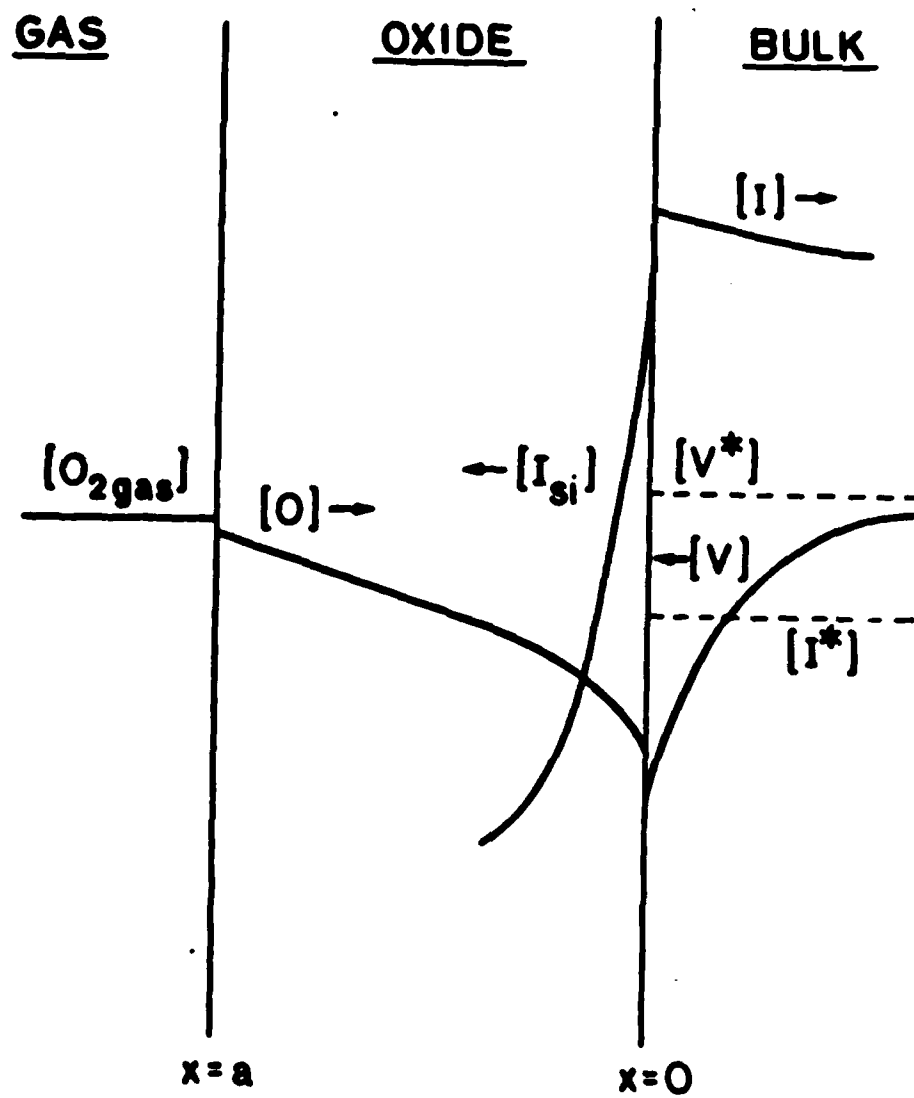[10.21] B. E. Deal, J. Electrochem. Soc., 110, 527 (1963).

Fig. 10.1: Densities and fluxes of reactants in oxidizing silicon.
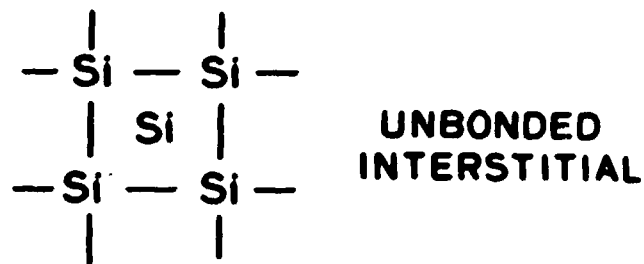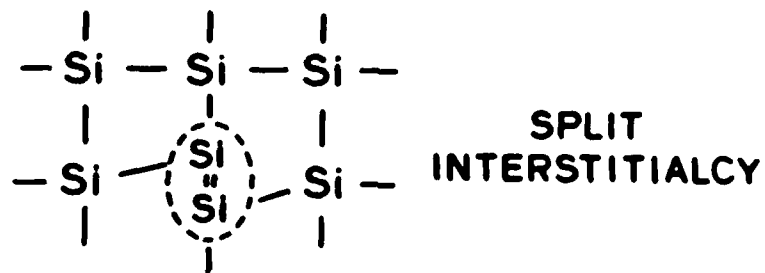
Fig. 10.2: Pure interstitial in silicon lattice.

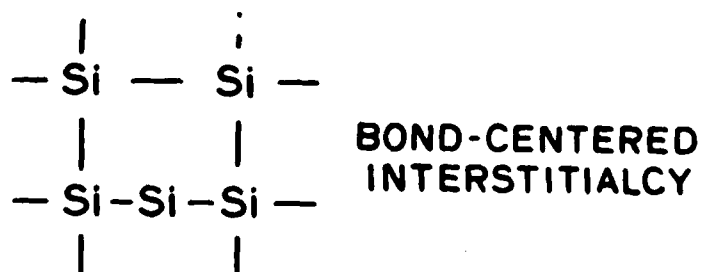Fig. 10.3: The split interstitialcy - two Si atoms sharing a single lattice site.

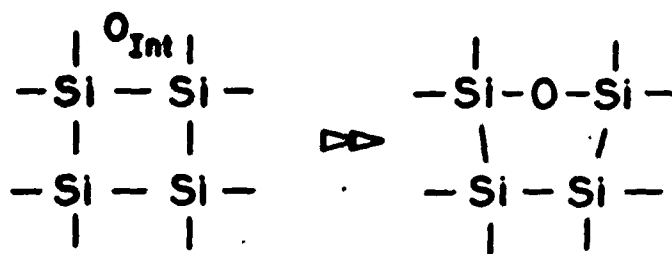Fig. 10.4: The bond centered interstitialcy - a second possible configurtion.
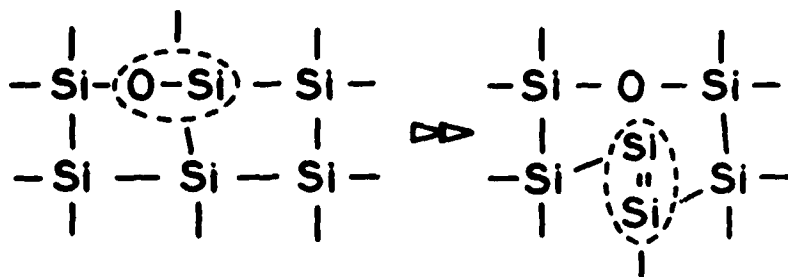
Fig. 10.5: Single oxidation reaction

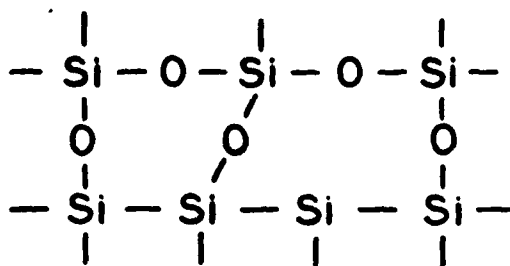Fig. 10.6: Creation of silicon interstitialcy at interface.

Fig. 10.7: Lattice fit at interface.

279

11. OXIDATION ENHANCEMENT TECHNIQUES AND THE PHYSICAL MODELING OF SILICON OXIDATION

W. A. Tiller, Y. T. Thathachari, W. E. Dibble, Jr.,

D. N. Modlin, and E. M. Young

11.1 INTRODUCTION

Although a great many investigators have probed the ideal and defect structure of silica both theoretically and experimentally, we still do not have a completely acceptable picture of its structure that allows us to correlate crystalline vs. vitreous features, stoichiometric vs. non-stoichiometric features, point defect vs. electron state features, interstitial vs. substitutional diffusion features, etc. In the present atomistic modeling of the thermal oxidation of silicon [11.1-11.4], the structure of the near interfacial region is of crucial concern because this is thought to be a key determining factor in its role as a diffusion blocking layer [11.4] and in the formation of both fixed oxide charge and interface charge states.

To shed more light on this picture, two theoretical and two experimental studies were instituted some time ago, and some of their results will be reported here: (1) steric origin of the Si-O-Si angle distribution in silica, (2) energetic interactions of foreign species with silica, (3) electric field-enhanced oxidation of silicon, and (4) laser light enhanced oxidation of silicon. A discussion of this work follows.

11.2 STERIC ORIGIN OF THE Si-O-Si ANGLE DISTRIBUTION IN SILICA

We initially partitioned the structural features into primary building block considerations (the SiO_4 tetrahedron), dimer unit considerations (Si_2O_7 species) and ring structure considerations. The basic structural unit in silica is the oxygen tetrahedron surrounding a silicon interstitial with a free energy of formation $\approx 98\%$ of the total free energy of formation for any of the silica

280

polymorphs [1.5]. Negligible deformation occurs in these tetrahedral units as they corner-share bridging oxygens to form the different polymorphs. Thus, a description of these various structures reduces to a description of the relative orientation of adjacent tetrahedra; i.e., the dimer unit.

We have found that the dimer units can be fully characterized in terms of the three angles θ, δ_1 and δ_2. The angle θ is made by two silicons at the bridging oxygen and is commonly called the TOT angle ($\approx 144°$ in quartz). The dihedral angles δ_1 and δ_2 give the degree of rotation of these two tetrahedra at fixed θ relative to a chosen reference frame. The dimers combine to form ring structures in three dimensions which become the major macroscopic structure determinant. It appears that dimer parameters leading to rings other than 6-membered (Si_6O_{18}) are unlikely to exhibit long range order. Our first significant success with this new modeling approach has been to show that all the crystalline polymorphs of SiO_2 can be generated form a remarkably small set of values of θ, δ_1 and δ_2.

Most recently, the constraints on closed n-ring assemblies (Si_nO_{3n}) have been assessed to identify those constraints leading to possible vitreous structures. Vitreous silica data suggests a distribution of TOT angles from 120° to 180° with a peak at $\approx 144°C$ [11.6,11.7]. The origin of the TOT angular distribution has been discussed by many investigators [11.8-11.12] and many factors have been suggested as determinants of this angle; i.e., the covalent energy of the Si-O bond, the coulombic energy, the non-bonded Si-Si interaction in the dimer, etc. While these factors are undoubtedly relevant to determining the correct TOT angle distribution, in our studies we have found the O-O separation to be a singularly important and restrictive criterion for determining the dimer configurations.

A detailed investigation of interatomic distances in all known silica polymorphs showed that the O-O separation was considerably larger than the van

281

der Waals, or non-bonded contact, distance of 2.9 Å. We decided to make the O-O distance a restrictive criterion in a Monte Carlo study of acceptable TOT angles for three dimensionally connected silica dimers and trimers. We will discuss the dimer study here.

In the dimer study, we assumed perfect tetrahedra with a Si-O bond length of 1.60 Å. We chose one of the two tetrahedra as a reference and fixed it in space as illustrated in Fig. 11.1. Here, the bridging oxygen, O_1, is chosen as the origin while O_1-O_2 define the x-axis. The x-y plane is defined by Si, along with O_1 and O_2. The Si_1 lies on the perpendicular bisector of O_1-O_2 which is parallel to the y-axis. The line O_3-O_4 is parallel to the z-axis. The position of the second tetrahedron, defined by O_5, O_6, O_7 and Si_2, is determined by three angles These angles are chosen as the Euler angles, α, β and γ.

The second tetrahedron can be derived from the first by (i) rotation about the x-zxis by angle α (frame x, y', z'), (ii) rotation by angle β about the new z' axis (frame x', y'', z') and (iii) rotation by angle γ about the new x' axis (frame x', y''', z''). From Fig. 11.1, it is easy to see that the angle β is the angle O_5-O_1-O_2. The angles γ and α can be considered as dihedral angle rotations around O_1-O_2 and about O_1-O_5, respectively, with the reference plane being the being the plane O_5-O_1-O_2. In our dimer study, the angles α, β, and γ are chosen as random values uniformly distributed in the range 0-360°. No symmetry relations for these angles were invoked because we wanted to avoid any possible bias.

For each of the dimers thus specified, the TOT angle was calculated from the Si-Si distance. The O-O and Si-O distances for atoms in different tetrahedra were then computed, and if any O-O distance was below our allowed minimum of 2.9 Å, the dimer was abandoned as not having passed our restrictive criterion. A new dimer was then chosen by another random triad of angles. If all the O-O distances were larger than 2.9 Å, the dimer was accepted and its maximum O-O

distance noted (saved for future reference). Both the sample pool, based on 2400 dimers, and the accepted dimer distributions of TOT are shown in Fig. 11.2.

When no restrictions whatsoever are placed on the three dimer angles, it can be shown that the TOT distribution follows a sine curve as observed in Fig. 11.2. The shaded area represents the distribution for the accepted dimers; i.e., those where the 0-0 distance is not less than 2.9 Å. This distribution has cutoffs near the 120° on the low end and 180° on the high end of the scale with a maximum at ≈145°, all in excellent agreement with the experimental observations for silica. This result is the remarkable consequence of a single stipulation on the minimum acceptable 0-0 separation (2.9 Å) in adjacent tetrahedra and is not based in any way on a specification of detailed energetic considerations.

For these distributions, we chose boxes of 10° width and thus expected the frequency of occurrence of the three dimer angles α, β and γ to be ≈67 on the average. As shown in Fig. 11.3, this expectation is met for the three angles and they are seen to be uniformly distributed over the entire 0-360°C range.

Our finding, simply stated, means that based upon purely geometrical considerations, when adjacent tetrahedra assume random orientations, the TOT angles have a distribution centered around 144° and extending to 110° and 180° on the sides. This result does not mean that, in a given structure, there should be at least a pair of oxygens in the adjacent tetrahedra of the dimer with a separation of ≈2.9 Å. The actual separation found in a particular structure and consequently the TOT angle may be determined by a number   ither criteria, among which are the long range forces that determine the closec   tructures plus the dimer and trimer configurations.

Since the observed minimum oxygen separation in silica is much larger than the 3.0 Å usually expected in molecular structures, one might wonder if there is another potential function operating to yield a larger 0-0 contact distance, say

≈3.4 A. However, scrutiny of the low quartz structure shows us that the minimum O-O separation in adjacent tetrahedra is not between units sharing a corner but between units linked by an intervening unit; i.e., a trimer. Although the minimum O-O distance is 3.4 A in the dimer, it is 3.30 A in the trimer. Thus, the minimum O-O separation in a particular dimer is probably associated with the connectedness of that dimer to its adjacent ring structures.

Considering the accepted dimers of Fig. 11.2 a step further, we sorted them in terms of the minimum O-O separation using boxes of width 0.2 A and found the TOT angle for each box; i.e., the average plus the standard deviation. These results are presented in Fig. 11.4. Here, the vertical dimension in the histrogram is the standard deviation of the TOT angle for the box. For example, with a minimum O-O separation of 3.3 to 3.5 A, the TOT angle is ≈140°. Clearly, the small values of the minimum O-O distance correlates with the smaller TOT angles. The observed minimum separation and the TOT angles in known polymorphs of silica are in reasonably good agreement with this curve.

We may now ask if the stipulation of a minimum O-O distance is qualitatively different from a stipulation that the Si-Si separation is not to be less than 3.0 A, say. Our answer is yes! While the former naturally leads to the observed distribution of the TOT angles, the minimum Si-Si separation fixes only the minimum TOT angle. Finally, our results are helpful in a specific manner in the modeling of amorphous silica. Instead of starting with a distribution of the TOT angle, we need only a minimum allowed O-O separation and a random distribtuion of the dimer angles. We can then use the dimers from the accepted pool as described above to build more extended structures.

11.3 ENERGETIC INTERACTION OF FOREIGN SPECIES WITH SILICA

We have used the CNDO/2 molecular orbital method to obtain the interaction energy between various SiO_2 symmetry fragments (dimers, trimers, n-rings,

etc.). The recent work of de Jong [11.12,11.13] showed that the minimum cluster of atoms necessary to model the valence state of the bridging oxygen and those of neighboring silicon atoms in silica glasses consists of two SiO_4 tetrahedra sharing a bridging oxygen, as illustrated in Fig. 11.1; i.e., a dimer unit. In order to account for the effects of the surrounding structure, de Jong [11.13] required that four criteria be satisfied: (1) all the orbital energies, ϵ_i, of the occupied molecular orbitals should be negative, (2) the total energy, E_T, should show a minimum in its variation with TOT angle, θ, and $\theta \approx 140°C$ to be consistant with the radial distribution function for silica glass, (3) the binding energy of the molecule should be negative or close to zero and (4) the bond overlap population and other relative measures of bond energy should be consistent with the geometries and dimensions obtained from silicate structures. de Jong [11.13] found that the molecule $H_6Si_2O_7$ fulfills these conditions.

By placing hydrogen atoms on the non-bridging oxygens in $H_6Si_2O_7$, the electron density on the molecule could be decreased, and by suitable choice of the orbital exponent, bonding parameters and orbital electronegativity for hydrogen, identical overlap calculations for all Si-O bonds in the cluster resulted and identical charges on all the oxygen atoms were found. This is what one would expect for an average linkage in a completely polymerized silicate melt.

The fundamental chemical basis of our modeling using CNDO/2 is that chemical reactivity is determined solely by the electronic structure of the clusters of atoms used in the calculations. de Jong and Brown [11.12] calculated several electronic spectra for $H_6Si_2O_7$ using CNDO/2 and showed them to be nearly identical with the experimentally derived spectra for vitreous silica. Thus, because of the close similarity of electronic structure, it is felt that the $H_6Si_2O_7$ cluster will serve quite well to model chemical reactions between the bridging oxygens and various interstitial species.

Our approach to the present study is to start at the simplest silica structural element, the dimer, and proceed through the trimer to the various n-rings lying almost in a plane and finally to the most complex structural element, a series of n-rings connected three-dimensionally. Our previous computer program, inherited from de Jong [11.13], only allowed the handling of 82 orbitals so that a tetramer ring unit in a pseudoplane was the largest structural element that could be treated. This program is being expanded to handle 256 orbitals, but the present results are with the old program. Some additional limitations of the present method are that (i) the Si-O bond lengths are not optimized for minimum total energies, (ii) the reaction energies are determined only at 0°K and (iii) these calculated energies are not absolute so that only relative energies between different species or conditions are likely to be relevant.

11.3.A   Interaction Energies with the Dimer Unit

Using the $H_6Si_2O_7$ molecule in the eclipsed conformation, a neutral or charged species, X, is placed on the bridging oxygen in the upward direction at distance d from O(br) and lying in the Si-O-Si plane as illustrated in Fig. 11.5a. The total energy, E_T, is calculated. The reaction equation to evaluate the total energy change, ΔE_T, is of the form

$$H + H_6Si_2O_7 \rightleftarrows H_6Si_2O_7X^a; \quad \Delta E_T = \sum_P E_T^P - \sum_R E_T^R \tag{11.1}$$

where a, P and R refer to adsorbed, product and reactant, respectively. To date, we have obtained the data presented in Table I for the parallel $||$) and perpendicular ($\perp$) configurations of the molecules; i.e., $||$ or $\perp$ to the Si-O-Si plane.

The minimum energy geometries were initially assumed to be those of the highest symmetry. However, additional calculations on the reactions between O_2 and O(br) in the dimer have shown that the total energy minimum actually occurs

with the O-O axis bisecting the TOT angle but with one oxygen in the Si-O-Si plane while the other is out of the plane. Further calculations are underway to determine the precise minimum in E_T as a function of the rotation of the diatomic molecule ad its distance from O(br).

11.3.B Intraction Energies with the Tetramer Unit

Using the $H_8Si_4O_{12}$ molecule in the tetramer form, a variety of species X were placed in the ring at different coordinate positions, as indicated in Fig. 11.5b. The reaction energy was minimized for $X \equiv O$ located at the coordinates (0.5, 0, 0) with values of ΔE_T given in Table II for this and other choices of X. In evaluating the energies for charged species, the electron work function for SiO_2 was taken as 5.1 eV. From Tables I and II, we note that Si_I is an unfavored species where O_I is strongly favored, and we note that the H_2O interaction is stronger than the O_2 interaction which is consistent with the experimental solubility data for SiO_2.

11.3.C Formation of Vacancies and Frenkel Defects

CNDO/2 calculations using normal bonding parameters yield equilibrium Si-O distances, d(Si-O), greater than the values found in natural silicates. This led to problems in calculating E_T for V_O. To decrease the calculated equilibrium values of d(Si-O), it was necessary to increase the orbital exponent of Si, ξ_{Si}, from its normal value of 1.383 [11.13] to 1.86. This contracts the Si orbitals and produces a minimum in the total cluster energy at d(Si-O) =1.61. Using both dimers, D, and tetramer, T, as the reactant species, Table III provides the values of ΔE_T for various product species. Most important, we see that the decomposition of two dimer units at the silica surface to produce two V_O plus an O_2 gas molecule is strongly favored as is the reaction to form $V_O + O_I$. In addition, Frenkel defect formation in the bulk silica due to the reaction of 4 tetramers; i.e.,

$$4T \rightleftharpoons 2 \, (T\text{-}V_0) + 2(T\text{-}O_I); \quad \Delta E_T = +3.32 \text{ eV} \tag{11.2}$$

is also well favored with a Frenkel defect formation energy of 1.61 eV. This is a likely candidate for determining the bulk V_0 and O_I population for SiO_2. The values of ΔE_T for these reactions are given in Table III.

11.3.D Interaction Energies with Various Trimer Units

A variety of trimer structures, such as illustrated in Fig. 11.5c, have been considered. Comparing total energies of clusters having the same condensation (condensation $\equiv$ numer of bridging oxygens/number of non-bridging oxygens) but different configurations, one finds virtually no energy difference between them ($\Delta E \approx 100$ cal/mole). The large number of naturally occurring polymeric silica species undoubtedly results from the minimization of the total free energy via a large configurational entropy.

When the silica condensation $(O(br)/O(u\text{-}br))$ is changed, however, larger energy differences between clusters are observed. An interesting relationship between Si tetrahedra of differing condensation has been evaluated. For example, the total energy of an isolated Si tetrahedron with no $O(br)$, Q_0, is -78.382 atomic units (a.u.) for the standard parameterization. The situation for Q_1 (one(br) per tetrahedron), Q_2 and Q_3 are respectively -69.008 a.u., -59.089 a.u., and -49.170 a.u. Thus, the energy per bridging oxygen bond created is about +10 a.m.

11.4 ELECTRIC FIELD ENHANCED OXIDATION

For our negative point experiments, the Si wafer rested upon a quartz plate at a distance of 0.5 cm below the platinum needle indicated in Fig. 11.6 and a negative voltage was applied to the needle. Some results are shown in Fig. 11.7 which give oxide thickness profiles on Si <111> substrates for oxidation times of 0.25, 0.5, 1.0, 2.0 and 4.0 hours at 900°C in pure O_2. The needle current was held at -5μA during these experiments.

The bell-shaped profiles of oxide thickness mimic those of the ion beam intensity such that the peak in oxide thickness lies directly below the tip of the point electrode and the constant background level occurs where the beam intensity has fallen to zero. The peak oxide thickness is plotted as a function of time in Fig. 11.8 for this 900°C data. Here, each data point is the average of several identical experiments.

Both Si <111> and Si <100> data exhibit similarly cancave-upward curves with the <100> lying below the <111> for short times and being almost identical for long times. These results can be readily explained by noting that, at short times, the growth curves approach that of the native oxidation process while, at long times, the growth rate is dominated by the ion flux and approaches a unit slope on the log-log scale. If we make the assumption that the flux of oxidant due to the ion beam is independent of that due to the native growth process and that linear superposition holds, we expect that

$$X_B(r) = X_N + X_E(r) \tag{11.3}$$

where B, N, E and X refer to beam, native oxide, enhancement and thickness, respectively, while r refers to the distance coordinate from the beam axis. Since X_E is proportional to t while X_N is proportional to $t^{1/2}$, the former dominates at long t while the latter dominates at short t.

One of the most striking features of the negative point results is the relative insensitivity of the oxide thickness profiles to changes in temperature (see Fig. 11.9). As the temperature is increased, the peak oxide thickness, $X_B(0)$, remains relatively constant at 875 A for I = -5μA when one corrects for ion beam flicker which disappears at ~700°C. At T = 700°C, $X_B(0)$ begins to increase reaching 1200 A at 1000°C. This change is mainly due to changes in X_N with T because we expect that

$$X_E(0) = J_p t/qN_o \tag{11.4}$$

where J_p is the peak current density, t is oxidation time, q is the magnitude of the electronic charge and N_o is the density of O atoms in SiO_2.

The value of J_p is measured in a room temperature apparatus and is assumed to be independent of T, although first order analysis predicts it to decrease slightly with increase of T. It will also decrease slightly with increase of I because of charge repulsion effects. Assuming a value of $X_N \approx 20$ Å at low temperatures, one calculates $X_B(0) \approx 820$ Å for I = -5µA, t = 1 hr, and N_o = 4.5 x 10^{22} atoms/cm^3. This agrees fairly well with the value of 875 Å extrapolated from Fig. 11.9.

As T increases above 700°C, X_N contributes more strongly to X_B and, at 950°C with I = -5µA, X_N accounts for a thickness increase ≈30% over the due solely from $X_E(0)$. At 700°C, for t = 0.5 hr, $X_N \approx 40$ Å, as seen in Fig. 11.10. At these low temperatures, the ion beam flux essentially dominates the growth process.

11.5 LASER LIGHT ENHANCED OXIDATION

Over the course of this past year, our work in the area of photon-enhanced oxidation of Si has uncovered several interesting results. One experiment utilized a Lexel argon-ion continuous laser to generate a photon flux of selectable photon energy and adjustable intensity. Only one wavelength of the five most powerful wavelengths (2.4 eV, 2.5 eV, 2.54 eV, 2.6 eV and 2.7 eV) is utilized at a time. The beam was directed into a thermal oxidation furnace so as to hit the center surface of a vertical sample of Si at a few degrees off of the surface normal (see Fig. 11.11). This slight angle avoids multiple-reflection measurement errors in beam power density. Except for the presence of the beam spot, the silicon sample was otherwise exposed to typical thermal oxidation processing.

Typical oxidation conditions were 2 hours at 900°C in dry O_2 which resulted in control oxides of ≈ 450 Å and ≈ 300 Å for the <111> and <100> orientations,

290

respectively. The peak oxidaton enhancement (measured via ellipsometry) varied linearly with beam power density and was in the 3% to 30% range. Experiments were also carried out at low but constant beam power density, while the photon energy was varied. In this case, the percent oxidation enhancement was observed to increase linearly with beam wavelength. Thus, the enhancement is largely a photonic effect rather than a thermal effect.

REFERENCES

[11.1] W. A. Tiller, *J. Electrochem. Soc.*, 127, 3, 619 (1980).

[11.2] W. A. Tiller, *J. Electrochem. Soc.*, 127, 3, 625 (1980).

[11.3] W. A. Tiller, *J. Electrochem. Soc.*, 128, 3, 689 (1981).

[11.4] W. A. Tiller, *J. Electrochem. Soc.*, to be published in 1982.

[11.5] W. K. Kelley, "Contributions to the Data on Theoretical Metallurgy XIII," Bulletin 584, Bureau of mines (U.S. Government Printing Office).

[11.6] W. H. Baur, *Acta Cryst. B33*, 2615 (1977).

[11.7] J. H. Konnert and D. E. Appleman, *Acta Cryst. B34*, 391 (1978).

[11.8] W. A. Harrison, in *Physics of SiO$_2$ and Its Interfaces*, S. T. Pantelides, ed., Pergamon Press, NY, 1978, 105.

[11.9] C. Glidewell, *Inorg. Nucl. Chem. Lett.*, 13, 65 (1977).

[11.10] M. O'Keefe and B. G. Hyde, *Acta Cryst. B32*, 2923 (1976).

[11.11] G. V. Gibbs, M. M. Hammil, S. J. Louisnathan, L. S. Bartell and H. Yow, Am. Minerol. 57, 1578 (1972).

[11.12] B.H.W.S. de Jong and G. E. Brown, Jr., *Geochemica Acta*, 44, 1627 (1980).

[11.13] B.H.W.S. de Jong, "A Spectroscopic and Molecular Orbital Study on the Polymerization of Silicate and Aluminate Tetrahedra in Aluminosilicate Metals, Glasses and Aqueous Solutions," Ph.D. Thesis, Stanford University, December 1980.

Table I

X^a + $H_6Si_2O_7$ Eclipsed Dimer (TOT = 143.6°, ξ = 1.383)

Ad Species, X	ΔE_T (eV)
$F_2(||)$	-0.01301
$H_2(||)$	-0.03035
$O_2(||)$	-0.09106
$N_2(||)$	-0.28620
$H_2O(||)$	-0.39027
$C\ell_2(||)$	-0.41195
$OH^-(||)$	-0.73718
$HC\ell(||)$	-1.24453
$CO_2(||)$	-1.51771
H^+	-13.7462
O	-8.897
$O^=$	-0.6608
Si	+8.717
Si^{++}	+29.22
$O_2(\perp)$	-3.54

Table II

$$X^a + H_8Si_4O_{12} \text{ Tetramer Ring (TOT} = 143.6°)$$

Ad Species, X	Ring Position	ΔE_T (eV)
O	(0.50, 0, 0)	-3.04
$O^=$	(0.50, 0, 0)	+9.7
O^{++}	(0.50, 0, 0)	+9.0
O_2	(0, 0, 0)	+7.387
H_2O	(0, 0, 0)	+9.84
H^+	(0, 0, 0)	-10.28
OH^-	(0, 0, 0)	+1.88

Table III

Point Defect Formation Energies

Reaction Species	Product Species	ΔE_T (eV)
D	$D\text{-}V_0 + O(g)$	+8.97
2D	$2D\text{-}V_0 + O_2(g)$	+0.50
D^{++}	$D^{++}\text{-}V_0 + O(g)$	+0.28
D	$D^{++} + 2e^-$	+45.60
2D	$D\text{-}V_0 + D\text{-}O_I$	-1.07
2D	$D\text{-}V_0^{++} + D\text{-}O_I^{=}$	+21.75
3T	$2(TOV_0) + T\text{-}O_2\,(0,0,0)$	+7.88
$T + T\text{-}O_2\,(0,0,0)$	$2(T\text{-}O_I)$	-4.56
4T	$2(T\text{-}V_0) + 2(T\text{-}O_I)$	+3.32
T	$T\text{-}V_0 + O(g)$	+8.90

(a)

(b)

Fig. 11.1: (a) Tetrahedra dimers. (b) Coordinate system used. x,y axes are marked. z-axis is perpendicular to the plane of the paper. α, β, and γ are the three angles defining the dimer.
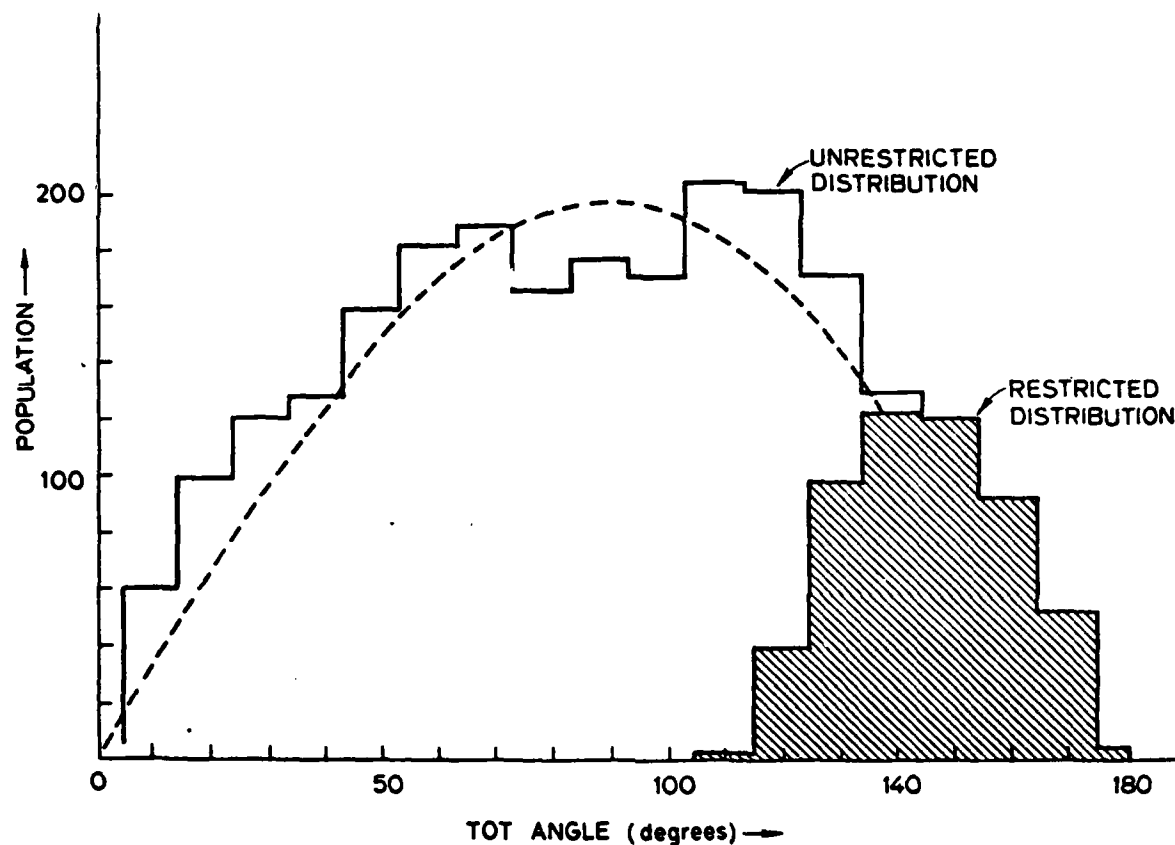
Fig. 11.2:   Distribution of the TOT angle: the x-axis is the TOT angle in
             degrees, the y-axis is the number of dimers with a specific TOT
             angle in boxes of width 10 degrees.   Shaded area represents the
             distribution in the dimers with oxygen-oxygen separation not less
             than 2.9 A.   The expected distribution of the TOT angle with no such
             restriction is a sine curve shown in dashed lines.   519 dimers were
             accepted out of a total of 2400 tried.
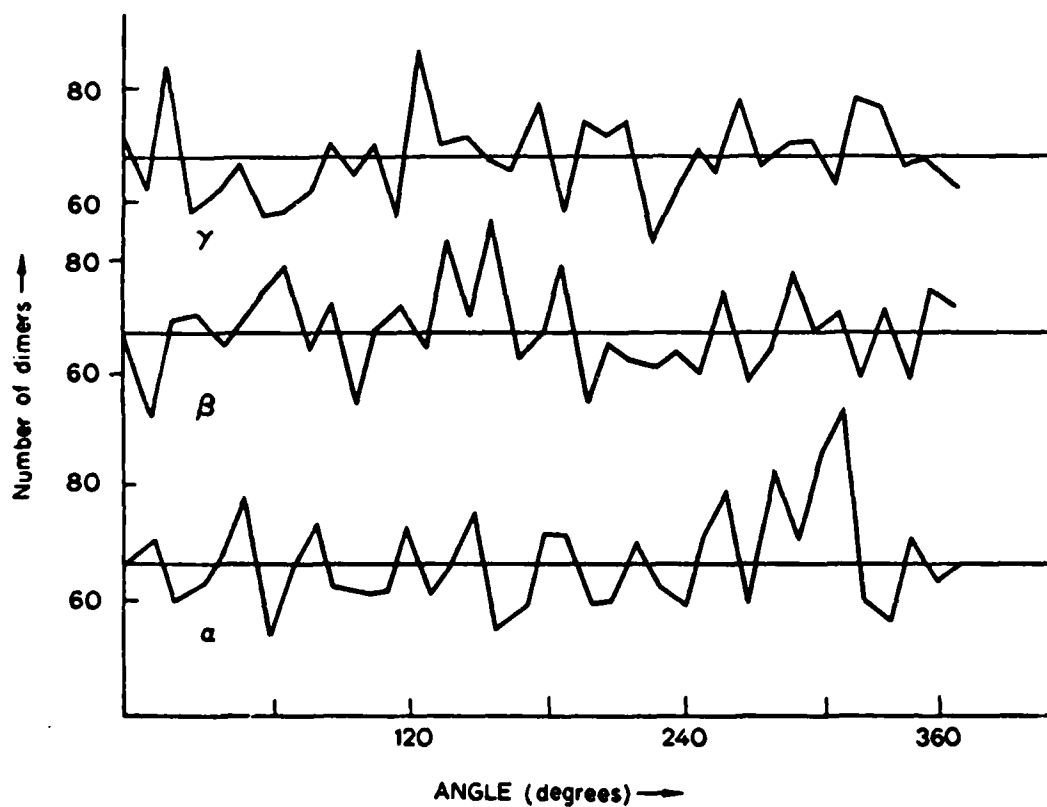
Fig. 11.3: The distribution of the three dimer angles α, β and γ: the x-axis is the angle in degrees, the y-axis is the number of dimers with the particular dimer angle in a box of width 10 degrees. The horizontal lines mark the average of 67, the expected value for the 36 boxes of width 10 degrees, for a total sample of 2400 dimers.
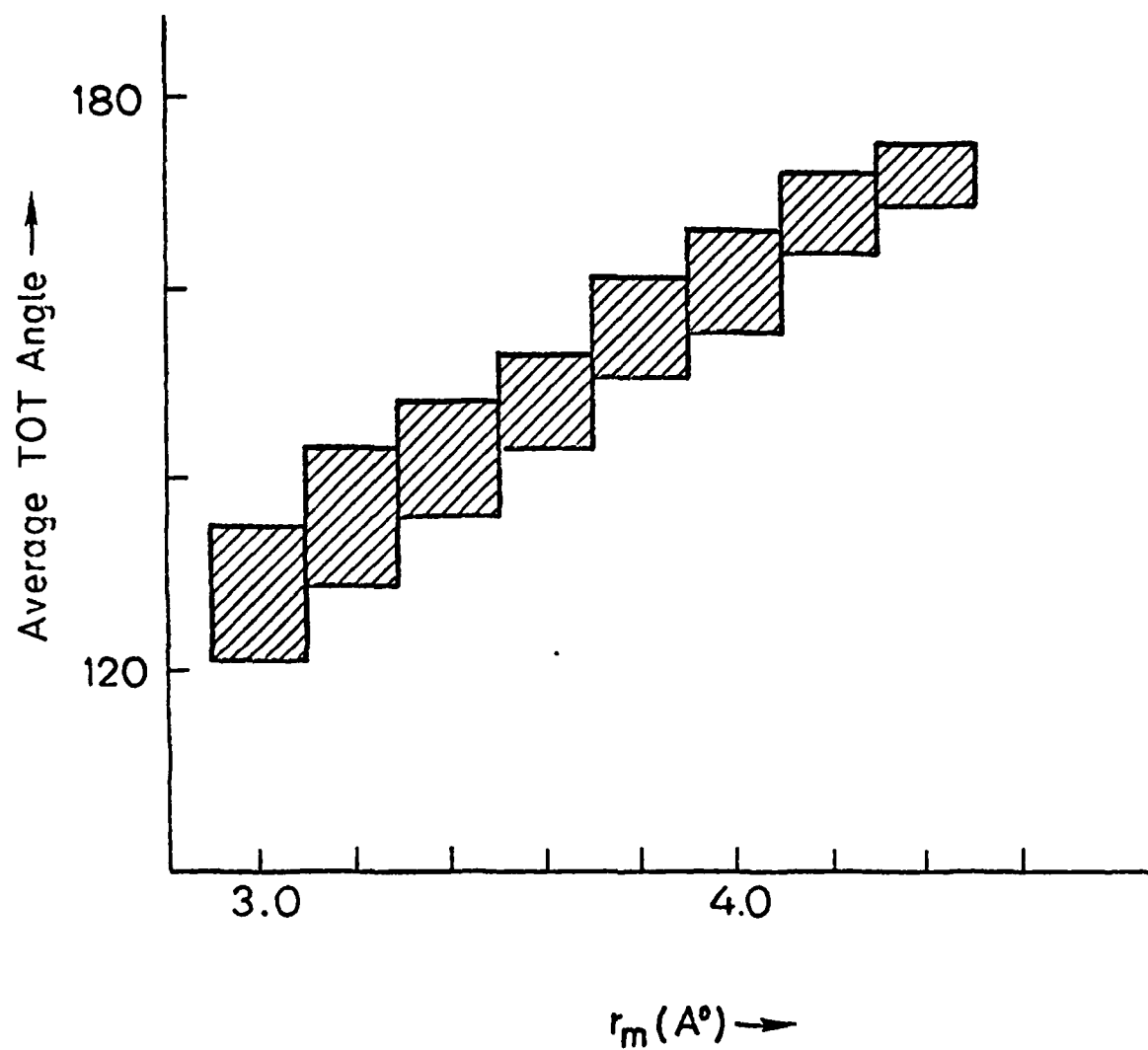
Fig. 11.4: The x-axis is the minimum 0-0 separation, rm, in the accepted dimers in boxes of width 0.2 A. The y-axis is the average TOT angles for the dimers in each such box, the vertical line representing the standard deviation in the angle.
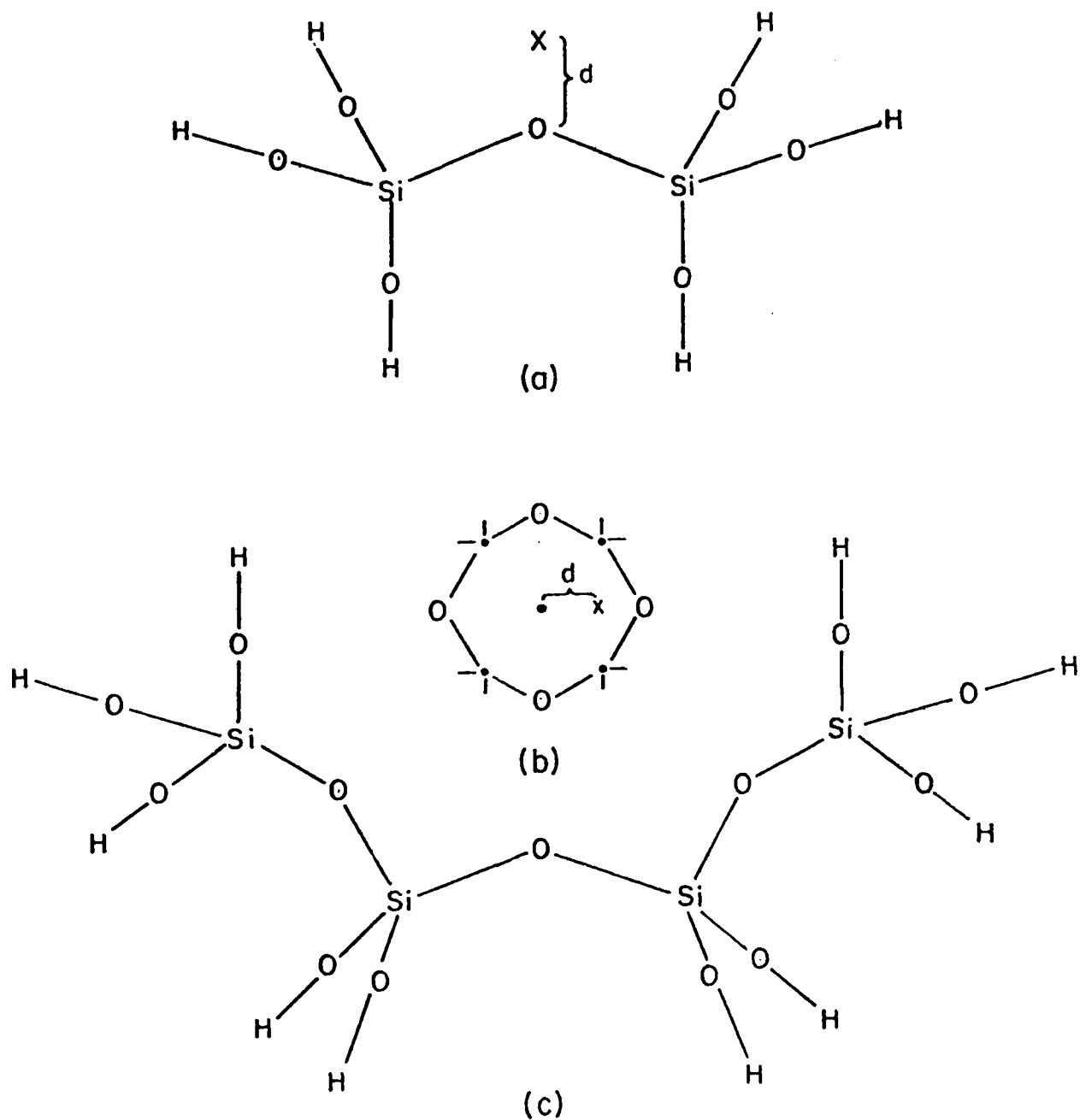
298

Fig. 11.5: Schematic illustrations of adsorption of species X to various silica fragments at distance d, (a) silica dimer, (b) silica tetramer ring, (c) silica trimer.

299

EXPERIMENTAL APPARATUS

QUARTZ FURNACE TUBE

PLATINUM
NEEDLE

GAS
INLET
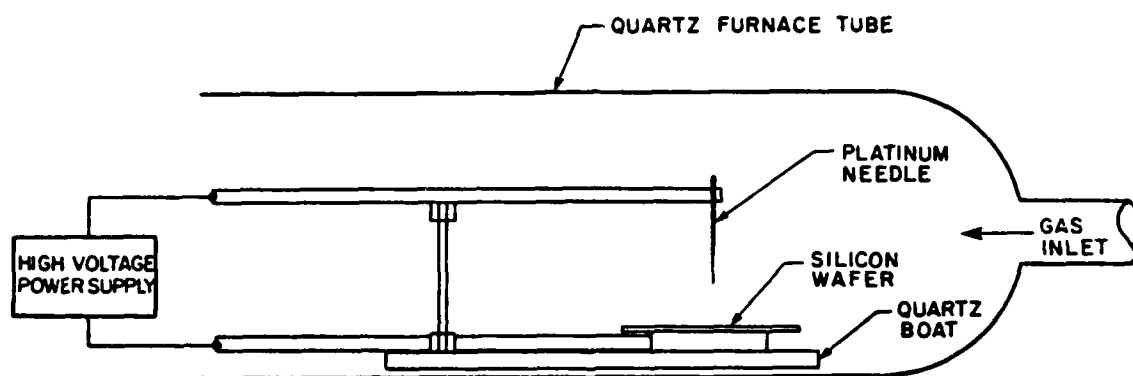
HIGH VOLTAGE
POWER SUPPLY

SILICON
WAFER

QUARTZ
BOAT

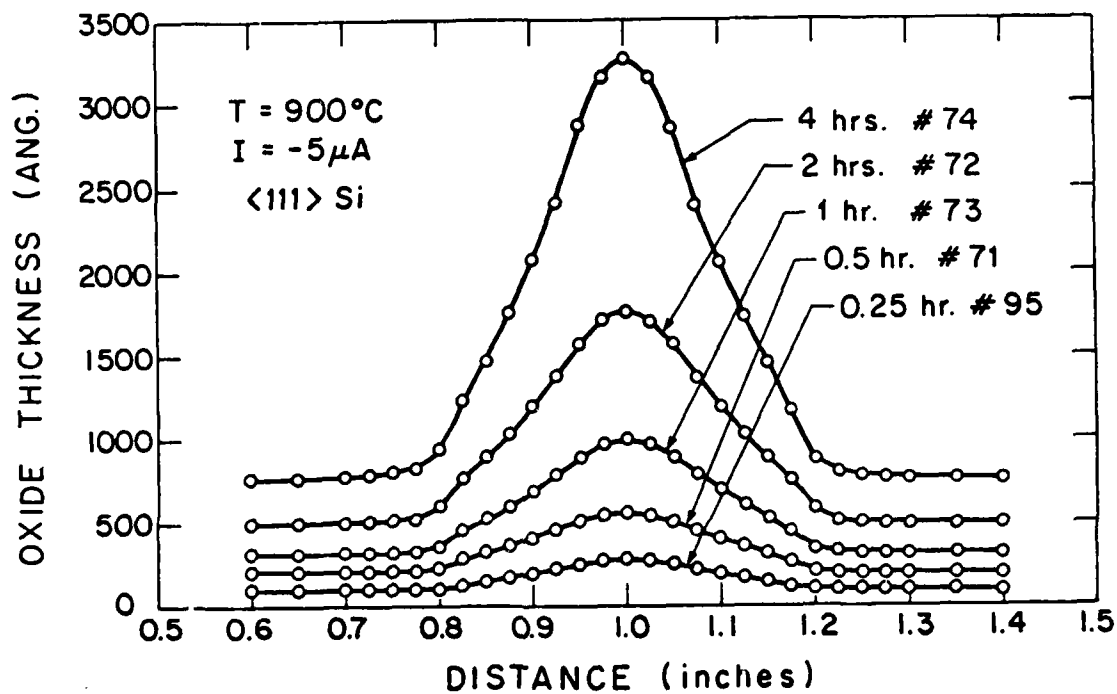Fig. 11.6: Schematic illustration of system for applying an electric field to an oxidizing Si wafer.

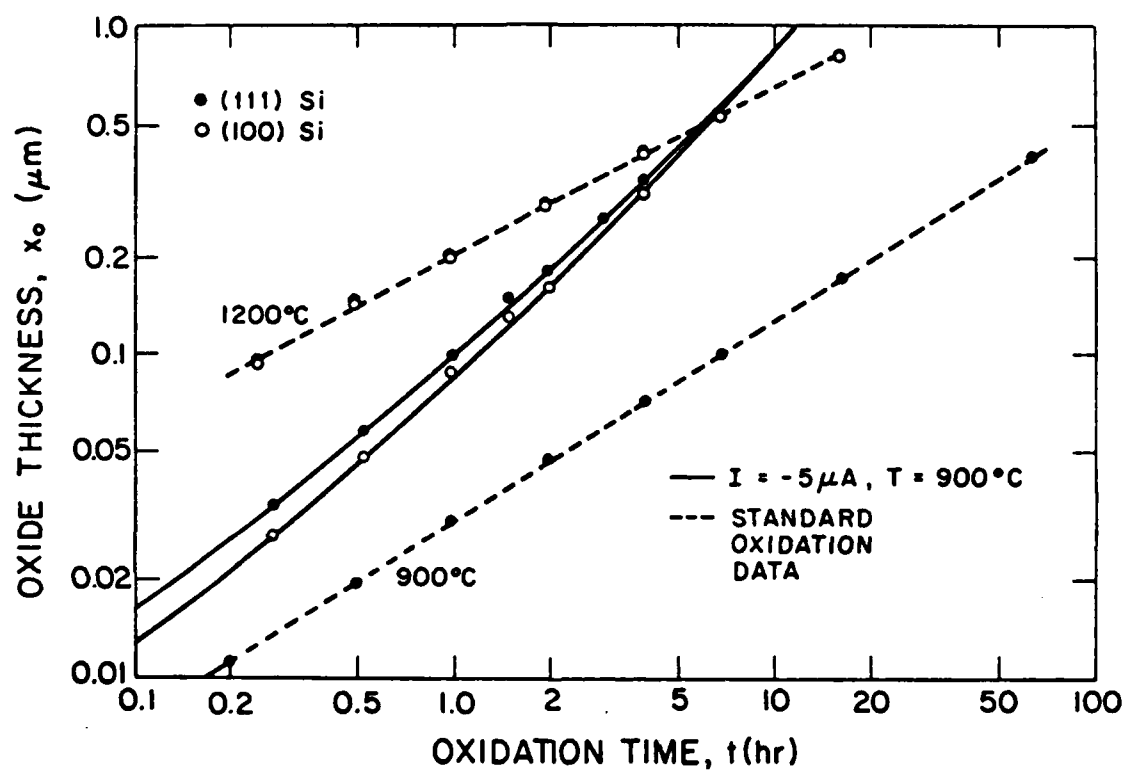Fig. 11.7: Oxide thickness profiles due to the ion beam at I = -5μA for a 900°C oxidation with a Si <111> substrate and various times.

Fig. 11.8:   Peak oxide thickness as a function of oxidation time at I = -5μA and
T = 900°C compared with the normal condition.

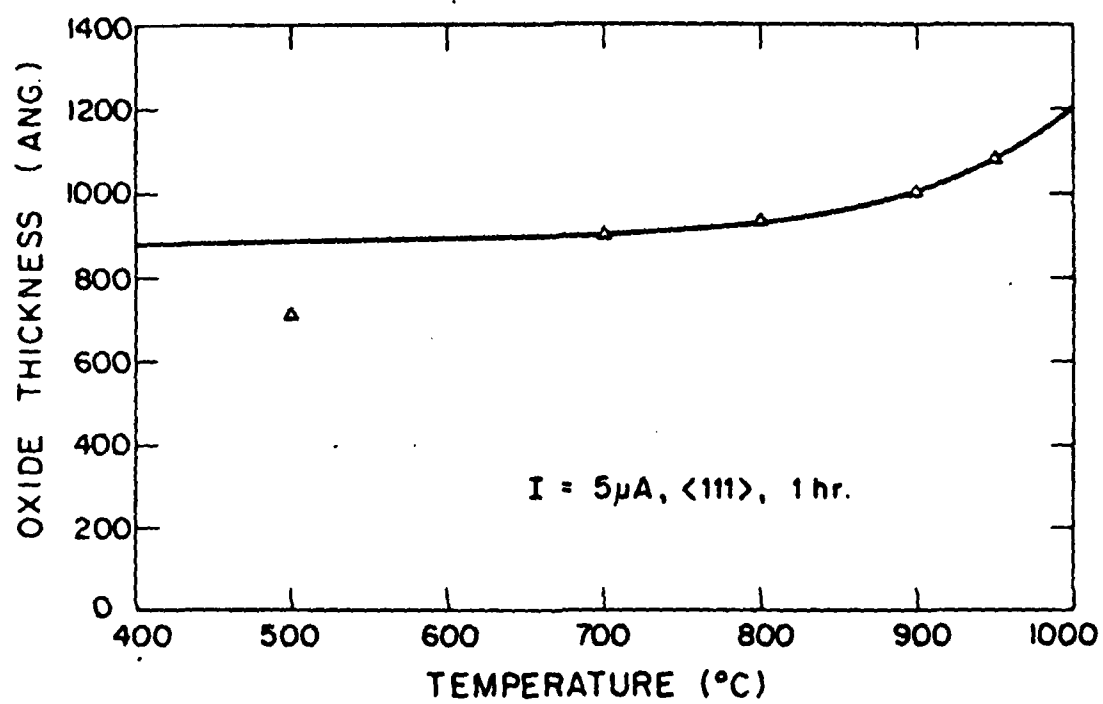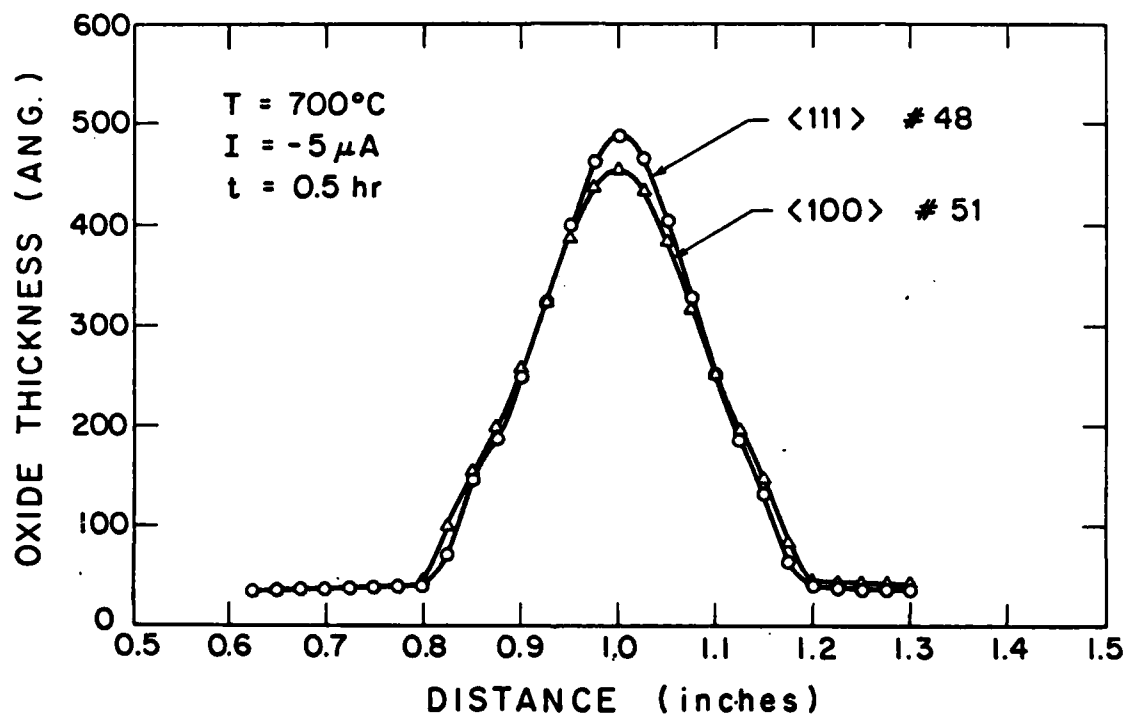Fig. 11.9: Peak oxide thickness as a function of temperature at $I = -5\mu A$ and t = 1 hr on a Si <111> substrate.

303

Fig. 11.10: Oxide thickness profile for <111> and <100> Si orientations at 700°C, I = -5µA and t = 0.5 hr.

RESISTANCE – COIL
FURNACE

SILICON
SAMPLE

SILICA
PEGS

HEAT–RETAINING
BAFFLE

LASER

(Beam diameter 1.5 mm)
(at surface of sample)
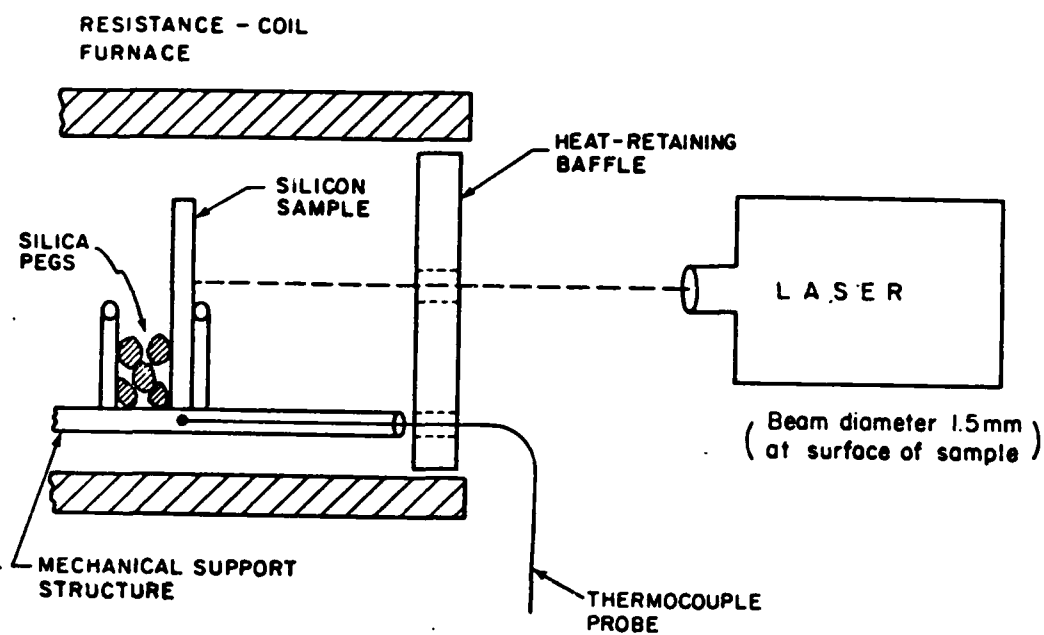
MECHANICAL SUPPORT
STRUCTURE

THERMOCOUPLE
PROBE

Fig. 11.11: Schematic diagram of the general apparatus used for the laser
stimulated oxidation studies.

12. ION IMPLANTATION RANGE STATISTICS

L. A. Christel, J. F. Gibbons

12.1 INTRODUCTION

Ion implantation has found extensive use in introducing controlled quantities of impurity-dopant atoms at desired depths into the target substrate in silicon integrated circuit technology. In this process, the impinging dopant ions enter the sample and deposit their energy in a complex series of scattering events, eventually coming to rest in a spatial distribution within the target.

The use of ion implantation in silicon integrated circuit fabrication has some distinct advantages over doping by high temperature diffusion processes. The total dose of impurities can be precisely controlled by accurate monitoring of the ion beam current and time of exposure. The maximum concentration of the implanted ion distribution may be placed at any desired depth by appropriate selection of the ion energy. It is even possible to customize the ion profile by performing multiple implants at various doses and energies.

The reproducibility and uniformity across the wafer, and from wafer to wafer, which are obtainable using ion implantation are essential for such applications as boron implants for threshold control in silicon-gate n-channel MOS transistors. Further, as such devices are scaled to smaller dimensions with gate oxide thicknesses approaching 200 A or less, the requisite shallow source-drain junctions are achievable and perhaps practical only with ion implantation.

As implantation applications grow in number and significance, detailed understanding and accurate modeling of the process and its applications correspondingly become more essential. Such modeling also can no longer be considered n a one-dimensional "isolated" sense. Rather, as device dimensions shrink both laterally and vertically, interactions of implantation with other processes

involved in IC technology may dominate such very small structures. Two-dimensional modeling of such interactions becomes vital and additionally will contribute enormously to understanding the basic physical mechanisms of those other processes themselves.

The successful utilization of ion implantation depends on the ability to predict and control the electrical effects which result from given implantation conditions. To this end, the ion stopping theory of Lindhard, Scharff and Schiott (LSS) [12.1-12.5] has proven invaluable as a means of approximately predicting primary ion range and damage distributions in semi-infinite targets. The LSS theory is used to calculate spatial moments (i.e. the mean and standard deviation) of the distributions, from which actual profiles must be constructed. Unfortunately, the LSS approach suffers from several distinct drawbacks. Specific areas of the implantation process itself as well as its implications and interactions have not been well understood nor adequately modeled.

As in the threshold control application already noted, practical device processing very frequently involves implantation into target structures composed of a substrate on which one or more thin films of varying composition have been grown or deposited. LSS theory assumes an isotropic or amorphous semi-infinite target. The effects of such a multi-layered target on the ion range distribution are not considered. Also not modeled is the experimentally observed dopant profile's deep channeling tail that results from ions scattered into the relatively open "channeling" directions of the crystalline silicon substrate. Furthermore, "knock-on" atoms from a masking surface layer which are recoiled into the substrate from collisions with the primary ion beam can critically affect electrical properties and device performance. This "recoil implantation" process also cannot be analyzed by LSS theory.

The substrate itself can be severely damaged by implantation. An initially crystalline substrate may be driven amorphous or large densities of vacancy-

307

interstitial point defect pairs created. Such damage distributions are integrally involved in implant annealing activation, but have not been adequately modeled. An approach similar to that of LSS has been used by Winterbon et. al. [12.6] to calculate energy deposition profiles. In this case however, the assymetry of such profiles necessitates the use of many moments for an accurate description. In addition, non-nuclear effects including electronic stopping must be neglected.Brice [12.7] has circumvented these problems to some extent by using a variation of the LSS moments approach to determine the more symmetric intermediate range distributions, and then calculating energy deposition profiles directly from these. Unfortunately, such approaches again require the assumption of a homogenous target and therefore fail to address the problem of multilayered media.

As device geometries diminish and new structures develop, understanding and modeling of all these aspects of implantation, as in SUPREM, will aid immeasurably in optimizing processes and minimizing costly and time-consuming empiricism.

12.2 BOLTZMANN TRANSPORT EQUATION APPROACH TO RANGE STATISTICS

During the past several years, emphasis in this project has been oriented very strongly toward modeling of the detailed mechanisms of the implantation process itself. A new approach for the direct calculation of primary ion and recoil range and damage distributions was one major development. By numerical integration of the linearized Boltzmann transport equation, ion range and damage distributions in multilayered targets, including recoil effects, have been successfully modeled [12.8].

In this transport equation (TE) approach, quantities of interest are determined by calculating in an step-wise 'fashion the momentum distributions of both primary and recoil particles as a function of distance z from the target surface. It is asumed that the target is amorphous or aligned in a random direction and possesses translational symmetry in all directions parallel to its surface. Thus

only two components of momentum, or equivalently an energy and direction cosine, are necessary to describe a particle's state of motion. Let $F(\vec{p},z)d\vec{p}$ be the number of particles (integrated over time) with momentum in the interval $d\vec{p}$ about $\vec{p}$ at depth z. There will be one such distribution function for each projectile type. The spatial evolution of each distribution is governed by the Boltzmann transport equation:

$$\frac{\partial F(\vec{p})}{\partial z} = N \int \left\{ \frac{F(\vec{p})\,d\sigma(\vec{p}' \to \vec{p})}{\cos \theta_{\vec{p}'}} - \frac{F(\vec{p})\,d\sigma(\vec{p} \to \vec{p}')}{\cos \theta_{\vec{p}}} \right\} + Q(\vec{p}) \qquad (12.1)$$

Here for simplicity we assume a single element target with number density N. The total differential scattering cross section $d\sigma(\vec{p} \to \vec{p}')$ represents a differential area presented to an incident ion by each target atom. Any ion which enters this area with momentum $\vec{p}$ is scattered to momentum $\vec{p}'$. A particle which scatters to a momentum (or energy) below some fixed limit is considered stopped, and is removed from the distribution. In this way the range profiles of both primary and recoil particles are generated. The quantity Q is a generation term which (for recoils) allows particles to be created from rest.

The distribution functions are assumed known at the surface plane z = 0. Recoil distributions are identically zero there, and the momentum distribution of the primary ion is taken to be a delta function:

$$F(\vec{p},0) = \Phi\delta(\vec{p}-\vec{p}_0) \qquad (12.2)$$

where Φ is the total dose (per cm^2) and $\vec{p}_0$ is the (unique) momentum of ions in the incident beam. With these initial conditions, the coupled transport equations are integrated numerically to determine the distributions at all depths z. Crossing a material interface, as occurs in a multilayered target, necessitates

309

changing only the number density and cross sections to quantities appropriate to the new material and continuing the integration. No changes occur in the momentum distributions because these are continuous functions of z for any target composition. Such an approach is thus easily applied to targets of practical interest.

In the past several years, elaboration of the transport equation model and implementation in a computer program have remained a major thrust of the work. Some of the many possible types of calculations, including recoil range distributions and stoichiometric disturbances in compounds, have been described in last years annual report and in several related publications [12.9,12.10,12.11].

During the past year, specific areas of activity have been:

1. Calculations of the total number of oxygen atoms which are recoil implanted from SiO_2 films into silicon during through-oxide implantation.

2. Development of a displacement criterion which can be used in conjunction with calculations to predict the extent of amorphous layers produced in silicon during implantation.

3. Improved calculations for very light ion implantation.

4. Initial investigation of very low energy ion implantation.

12.3 OXYGEN RECOIL YIELDS

As described in last year's annual report, the TE approach can be used to calculate the spatial distribution of oxygen atoms which have been recoil implanted into a silicon substrate during through-oxide implants. In addition, a simple procedure was developed which allows one to approximate the oxygen recoil distribution for arbitrary ion energy and oxide thickness. The approximate concentration $C(x)$ at distance x from the interface is assumed to be of the form

$$C(x) = Ae^{-x/L} \tag{12.3}$$

where the parameters A and L are given by

$$A = \left(\frac{1000 \text{ A}}{t_{ox}}\right) \cdot \phi_0 \cdot 8.7 \cdot 10^4 \cdot \exp\left[\frac{- \ln\left(\frac{R_p}{t_{ox}}\right)^2}{0.51}\right] \quad [cm^{-3}] \qquad (12.4)$$

$$L = 3.75 \text{ } \forall E_0 \quad [\text{Å}] \qquad (12.5)$$

where

E_0 = incident ion energy (keV)

ϕ_0 = incident ion dose (Cm^{-2})

t_{ox} = oxide thickness (Å)

R_p = projected range of ion in SiO_2 (Å)

$$\forall = \frac{4M_1 M_2}{(M_1 + M_2)^2}$$

An example of a recoil oxygen distribution is given in Fig. 12.1 where we compare the TE calculation for 200 keV As implanted into 1000 Å SiO_2/Si with the experimental results of Magee et al. [12.12]. One can see from the figure that although the exponential approximation agrees well in the deepest part of the distribution, there is significant disagreement near the interface.

One can correct for this discrepancy if one considers the recoil yield, defineu as the ratio of the total number of recoil atoms in the substrate to the number of incident ions. Experimentally, the recoil yeild is obtained by integrating the recoil range distribution and dividing by the incident dose. Since the recoil distribution is a maximum at the interface, the integral of the distribution, and hence the recoil yield, will be dominated by those recoil atoms which cross the interface with the lowest energies. These recoil generally result not from collisions involving the primary ion, but from secondary and higher

order collisions involving recoils which have been generated earlier in the film. Such "cascade" recoils are more difficult to account for in calculations.

When calculating the recoil range distribution, the recoils considered are only those which are created with an energy above some threshold (typically 1/40 of the initial ion energy). The neglect of these recoils has a negligible effect on the deeply penetrating part of the recoil distribution, but critically affects the value obtained for the total number of recoil atoms crossing the interface, and hence the recoil yield. In order to account for these cascade recoils, a modification of the TE calculation was made. With this modification, any recoil which is created by a particle moving with energy greater than the threshold is included.

In Figs. 12.2-12.4 are presented the results of calculations for the oxygen recoil yield as a function of oxide thickness for boron, phosphorus and arsenic, respectively, each at two different energies. One expects the recoil yield to be a maximum when the interface occurs at the depth of maximum energy deposition, typically about 0.8 times the projected range of the incident ion. This behavior is confirmed by the fact that the projected ranges (indicated by arrows in the figure) are slightly greater than the oxide thickness corresponding to the maximum recoil yield.

The yields presented here are higher than would be obtained by integrating the approximate exponential profiles. The extra yield is almost entirely contained in the first few hundred angstroms of the substrate. Thus these results can be used to correct the exponential approximation if the additional yield is assumed to lie near the interface.

A reasonable procedure would be to add a second exponential profile characterized by decay length L' and prefactor A', such that the total yield Y satisfies $\phi Y = AL + A'L'$. The question then is how to determine A' or L' (Since Y is

312

known, only one of A' and L' must be determined). One possibility is to assume that L' is equal to the range of the highest energy recoil which is neglected in the calculation of the recoil range distribution. Consider again the example illustrated in Fig. 12.1. A TE calculation indicates that the total yield in this case is about 1.6, whereas the exp ential approximation implies a yield of only 0.6. The recoils which are originally neglected are those with energies less than about $E_0/40 = 200$ keV/40 = 5 keV. The projected range of 5 keV oxygen in silicon is about 100 A. If we set L' = 100 A, then to account for the excess yield we must have A' = 10^{21} cm^{-3}. The dashed line in Fig. 12.1 shows the exponential correction with A' = 10^{21} cm^{-3} and L' = 10^{-6} cm. It can be seen that when this high yield correction is added to the original approximation, the experimental results are reproduced very well.

12.4 DISPLACEMENT CRITERION FOR AMORPHIZATION OF SILICON

It has been known for some time [12.13,12.14] that the complete activation of implanted dopants requires anneal temperatures of about 600°C if the substrate has been rendered amorphous by the implantation. However, temperatures of about 1000°C are necessary to remove compensating point and line defects if the substrate has been heavily damaged but is not amorphous. Since lower process temperatures are almost always advantageous (e.g. to prevent wafer warpage and dopant redistribution), it is of interest to determine the conditions under which a crystalline semiconductor such as silicon is rendered amorphous by implantation.

Vook [12.15] found that curves of the dose necessary to render the surface of silicon amorphous as a function of temperature for boron, phosphorus and antimony ions, could be collapsed into a single curve if one considered not the dose itself, but the amount of energy density deposited into atomic processes. This energy densty was taken to be the dose multiplied by the quantity v/Rp where v is

313

the total energy deposited into atomic processes per incident ion [12.16] and Rp
the projected range of the incident ion [12.17]. Such a procedure leads to the
conclusion that when the amount of energy density deposited into elastic atomic
processes exceeds a certain threshold, a continuous amorphous layer is formed.
Experimental evidence reviewed by Gibbons [12.18] and Vook [12.15] indicates that
this critical energy density lies in the range of 6-10 · 10^{20} keV/cm^3 for sub-
strates which are held at low enough temperature during implantation to insure
that all defects produced remain as stable damage.

Although this is a useful criterion for the production of an amorphous lay-
er, such estimates suffer somewhat from the fact that they do not acount for the
spatial distribution of the energy deposited into atomic processes. Since the
time of Vook's analysis, both the calculations of Brice [12.16] and the TE calcu-
lations [12.19] have shown that the damage is not uniformly distributed, but is a
maximum at a depth slightly less than Rp. Thus the simple energy density model
cannot, e.g., predict the extent of a buried amorphous layer which is known to be
produced by implantation of light ions such as boron into low temperature sub-
strates. To make such predictions, a more detailed knowledge of the spatial dis-
tribution of damage is necessary. The damage density profiles calculated by the
TE method give such information. Alternatively, one might use TE calculations to
determine the number of lattice displacements which occur at each depth. Relat-
ing the critical energy density to the number of lattice displacements is more
difficult because of the complexity of the recoil phenomena involved.

With the TE calculations, it is possible to determine quite accurately the
fraction of the silicon lattice which is displaced at each depth during ion im-
plantation. In the calculations, the momentum distributions are used in conjunc-
tion with the scattering cross section to determine the number of silicon recoils
produced per unit volume at each depth. Dividing by the number density of

314

silicon (5 · 10^{22} cm^{-3}) yields the fractional displacement of the lattice at that depth. These results have been correlated with the extent of implantation-induced amorphous layers which have been measured experimentally by channeling and backscattering of MeV He particles.

In Fig. 12.5 we present the experimental results of North and Gibson [12.20] obtained by channeling and backscattering of 2 MeV particles, for the disorder distribution resulting from implantation of 150 keV boron in silicon held at liquid nitrogen temperature. The backscattering spectrum obtained with the sample aligned in the <110> direction is normalized channel by channel to a spectrum obtained with the sample oriented in a random (non-channeling) direction. Thus a value of 1.0 indicates complete randomness of the sample and this has been found to imply that the sample is amorphous at that depth. The position of the edges of the buried amorphous layer produced by the implantation are taken to be the points at which the normalized yield drops half the distance to the lowest yield outside the amorphous region (yields of about .62 for the left edge and .77 for the right edge). Thus the amorphous region in this case is observed to lie between 0.15 and 0.55 microns.

Shown with the experimental data in Fig. 12.5 is the result of a TE calculation for the fractional displacement of the lattice under the same implant conditions, assuming displacement energies E_d (the energy necessary to remove a silicon atom from its lattice site) of 10 eV and 20 eV. Comparison of the calculated curves with the experimental results indicates that the fractional displacement at both edges of the amorphous layer is about 8.5% for E_d = 20 eV and about 12.5% for E_d = 10 eV. Thus in this case, it appears that about 10% lattice displacement results in amorphous silicon.

The results of TE calculations indicate that the number of displaced atoms is proportional to the energy density deposited into atomic processes. In Fig. 12.6

we compare the TE calculation with the results of Brice [12.16] for the energy deposited into atomic processes under the same conditions as in Fig. 12.5. It can be seen that for this case, in which the ion range distribution is known to be highly skewed, the assumption by Brice of Gaussian intermediate range distributions results in significant differences when compared to the TE calculation in which the energy deposition is determined more directly. The dashed horizontal line in Fig. 12.6 indicates the level of damage necessary to create an amorphous layer using the energy denstiy criterion that 10^{21} keV/cm^3 results in amorphous silicon. Note that at a normalized depth of 1.2, the TE calculation yields better agreement between the observed and predicted edge of the amorphous layer. If one combines the energy density criterion with the 10% atomic displacement criterion, one concludes that an average of 200 eV of deposited energy into atomic processes results in one displacement, i.e., the average energy given to a silicon recoil is about 200 eV.

Similar results for implantations of silicon into room temperature silicon at 100 and 200 keV [12.19] indicate again that about 10% lattice displacement is necessary to induce the amorphous transition at low temperature. Thus it appears that the 10% displacement criterion is rather independent of ion mass and energy as long as the substrate is held at sufficiently low temperatures.

12.5 LIGHT ION IMPLANTATION

When the ion to substrate mass ratio becomes less than unity, the possibility exists for ions to scatter through angles greater than 90° in a single collision. Ions involved in such collisions are actually traveling back toward the surface of the sample after the scattering event. Since the TE method as currently implemented involves forward integration in depth, those backscattered ions can be accounted for only approximately. In addition, the recoil effects that are associated with these ions are neglected. All these effects are notice-

able for boron in silicon (M_1/M_2 = 0.39) and quite significant for boron in Gallium Arsenide (M_1/M_2 = 0.15).

Toward the end of this fiscal year, work was started on an extension of the TE method to allow more accurate determination of profiles in the very light ion case. The idea is to apply a "multiple pass" approach. At each step in the forward integration, the source density of new backscattered ions is determined as part of the ordinary calculation. Instead of assuming such ions are stopped, the source density is stored for later use. After the calculation has proceeded once through the target, a reverse calculation begins, starting at the deep end of the profile, and proceeding back toward the surface. At each step, the source function which was previously stored is recalled and added back into the momentum distribution. If a third pass is desired, another set of source densities is stored (these now are ions which have scattered back toward the surface once, but have reversed their motion again to be travelling forward). The calculation proceeds until the profiles converge to a final result.

In Fig. 12.7, we show the results of a multiple pass calculation for 50 keV boron implanted into GaAs. It can be seen that after four passes that the profile has converged to the final result. So far, the method has been implemented only for the primary ion. Work in the future will involve the implementation of this approach to recoil and damage effects.

12.6 SUMMARY AND PLANS FOR FUTURE WORK

The one-dimensional version of the Boltzmann transport equation approach is essentially complete. A simplified version, which calculates primary ion range distributions in multilayered targets has been implemented in SUPREM III. Recoil and damage effects may be added as the need and/or desirability arise.

In addition to those areas already discussed, the work of the next several years must necessarily involve research which anticipates the major trends now

occuring in the integrated circuits industry. As devices shrink to sub-micron dimensions, the thickness of overlying films, such as oxides, must correspondingly be reduced. This will in turn require implantation energies to be lowered. At present, there exists very little theoretical or experimental data for ion range and damage distributions at energies below 10 keV. It is expected that the Boltzmann transport equation can be extended to the calculation of range and damage distributions for ions that are implanted at very low energies.

In addition to studies in low energy implantation, the decreasing lateral dimensions of devices requires a more accurate knowledge of mask edge effects occurring during implantation. One would like to be able to predict the extent to which ions penetrate laterally beneath a masking layer. The extension of a Boltzmann type of approach to two dimensions is possible in principle, but the implementation of such an approach would require prohibitive amounts of computing time. Some initial work should be started, however, in an attempt to determine what options are available with regard to this type of approach. Either a simplified Boltzmann type approach or perhaps an entirely new method will eventually be necessary if a complete two-dimensional process modeling program is to become a reality.

REFERENCES

[12.1] J. Lindhard, M. Scharff and H. E. Schiott, Mat. Fys. Medd. K. Dan. Vidensk. Selsk. 33, No. 14 (1963).

[12.2] H. E. Schiott, Mat. Fys. Medd. K. Dan. Vidensk. Selsk. 35, No. 9 (1966).

[12.3] J. Lindhard and M. Scharff, Phys. Rev., 124, 128, (1961).

[12.4] J. Lindhard, Mat. Fys. Medd. K. Dan. Vidensk. Selsk., 34, No. 14, (1965).

[12.5] J. Lindhard, V. Neilsen and M. Scharff, Mat. Fys. Medd. K. Dan. Vidensk. Selsk., 36, No. 10, (1968).

[12.6] K. B. Winterbon, Rad. Effects, 13, 215 (1972).

[12.7] D. K. Brice, J. Appl. Phys., 46, 3385 (1975).

[12.8] L. A. Christel, J. F. Gibbons, and S. Mylroie, J. Appl. Phys., 51, 6176 (1980).

[12.9] L. A. Christel, J. F. Gibbons and S. Mylroie, Nuc. Inst. Meth., 182/183, 187 (1981).

[12.10] L. A. Christel, and J. F. Gibbons, J. Appl. Phys., 52, 4600 (1981).

[12.11] L. A. Christel and J. F. Gibbons, J. Appl. Phys., 52, 5050 (1981).

[12.12] T. J. Magee, C. Leung, H. Kawayoshi, L. J. Palkuti, B. K. Furman, C. A. Evans, Jr., L. A. Christel, J. F. Gibbons and D. S. Day, Appl. Phys. Lett.,39, 564 (1981).

[12.13] B. L. Crowder and F. F. Morehead, Appl. Phys. Lett., 14, 313 (1969).

[12.14] B. L. Crowder, J. Electrochem. Soc., 117, 671 (1970).

[12.15] F. L. Vook, Radiation Damage and Defects in Semiconductors, (London Institute of Physics, 1972), p. 60.

[12.16] D. K. Brice, Ion Implantation Range and Energy Deposition Distributions, Vol. 1, (Plenum Press, New York, 1975).

[12.17] J. F. Gibbons, W. S. Johnson and S. Mylroie, Projected Range Statistics in Semiconductors, 2nd Edition, (John Wiley and Sons, 1975).

[12.18] J. F. Gibbons, Proceedings of the IEEE, 60, 1062 (1972).

[12.19] L. A. Christel, J. F. Gibbons and T. W. Sigmon, J. Appl. Phys., 52, 7143 (1981).

[12.20] J. C. North and W. M. Gibson in Ion Implantation, edited by F. H. Eisen and L. T. Chadderton (Gordon and Breach, London 1971) p. 143.
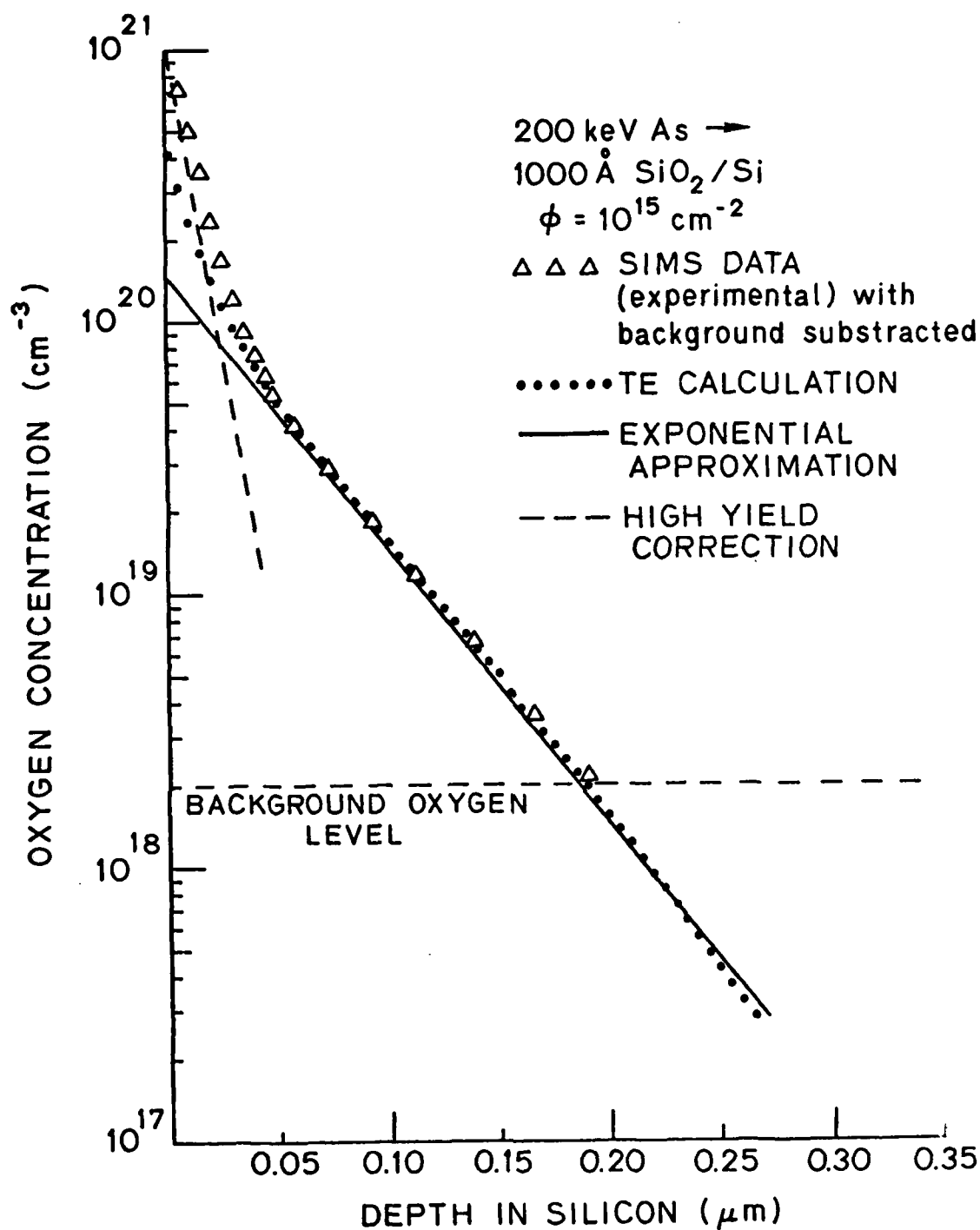
Fig. 12.1: Comparison of a TE calculation with experimental results from Magee et. al. [12.12] for the oxygen recoil distribution resulting from implantation of 200 keV As into 1000 Å SiO_2 on Si to a dose of 10^{15} cm^{-2}. Also shown is the analytic (exponential) approximation (solid line) and the high yield correction (dashed line).
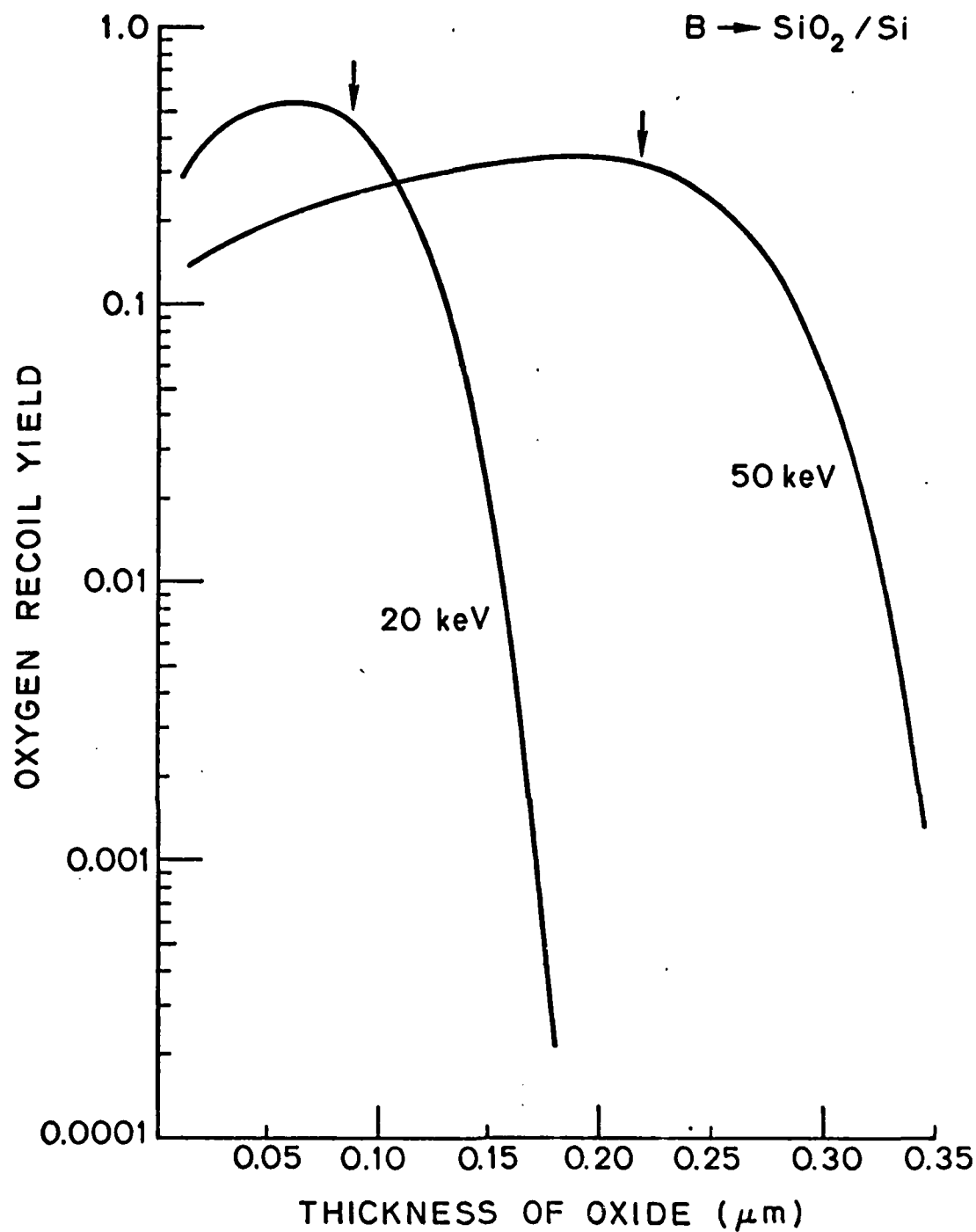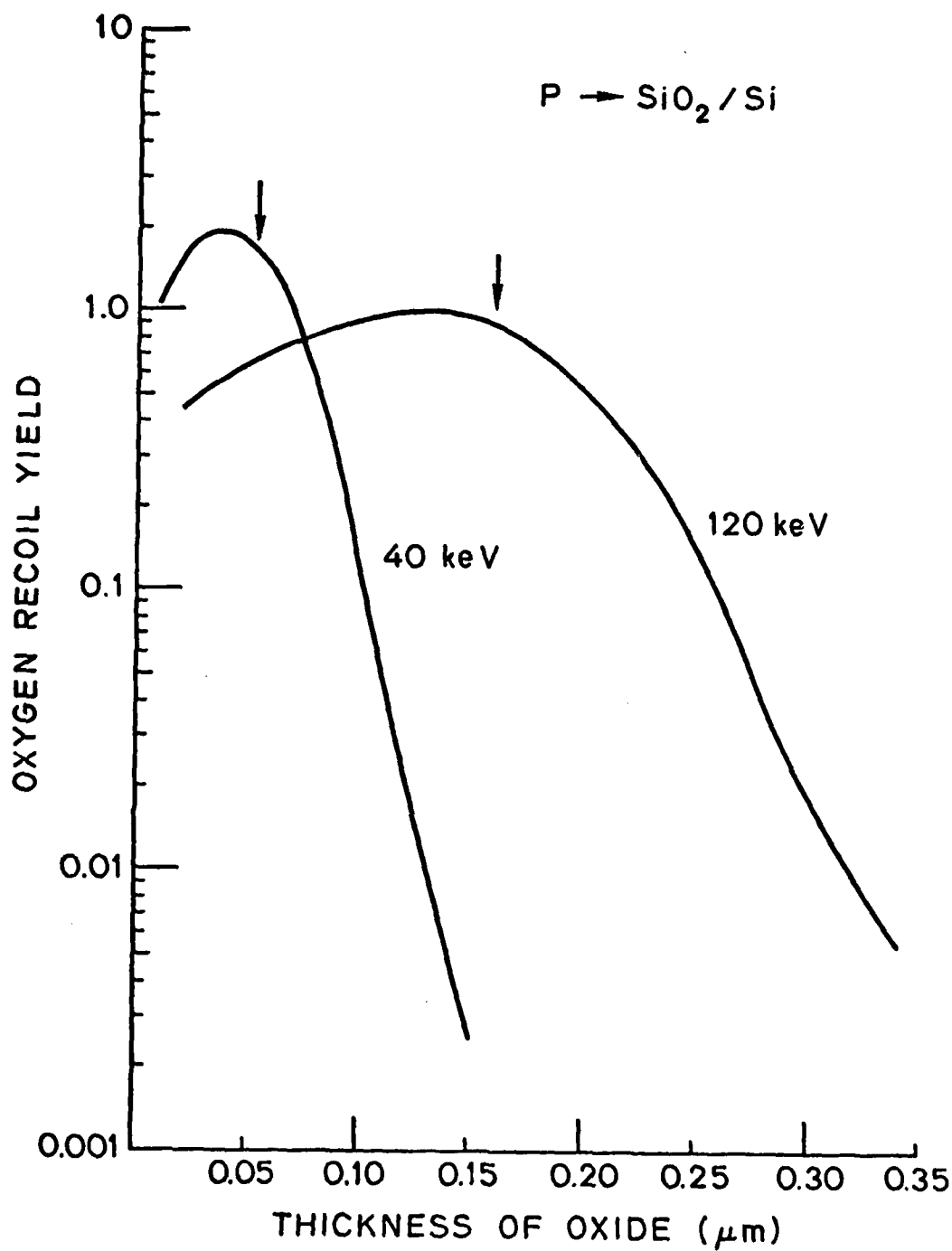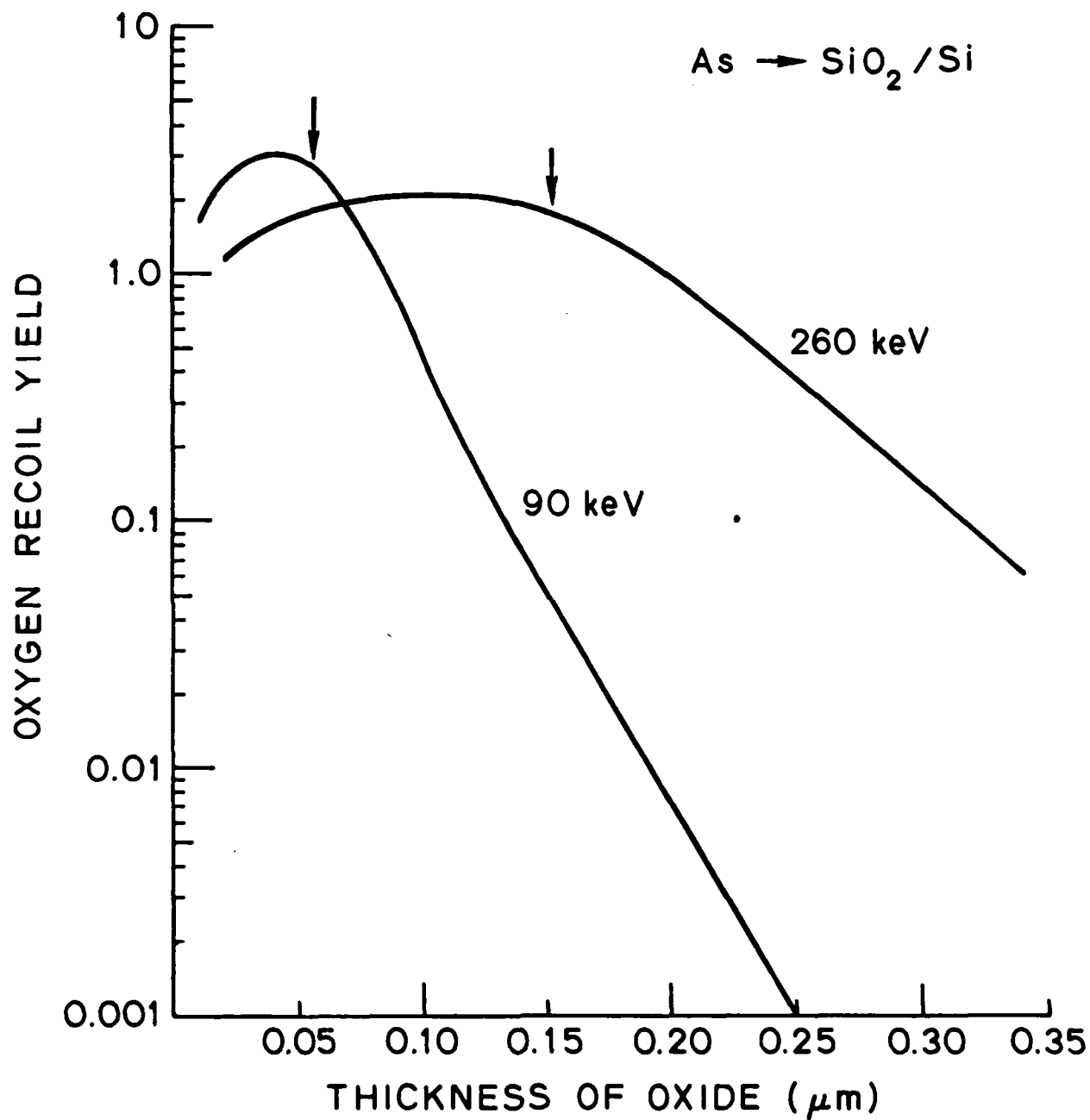
Fig. 12.2: Oxygen recoil yield as a function of oxide thickness for implantations of boron into SiO_2 on silicon. Arrows indicate the projected range R_p for each energy.
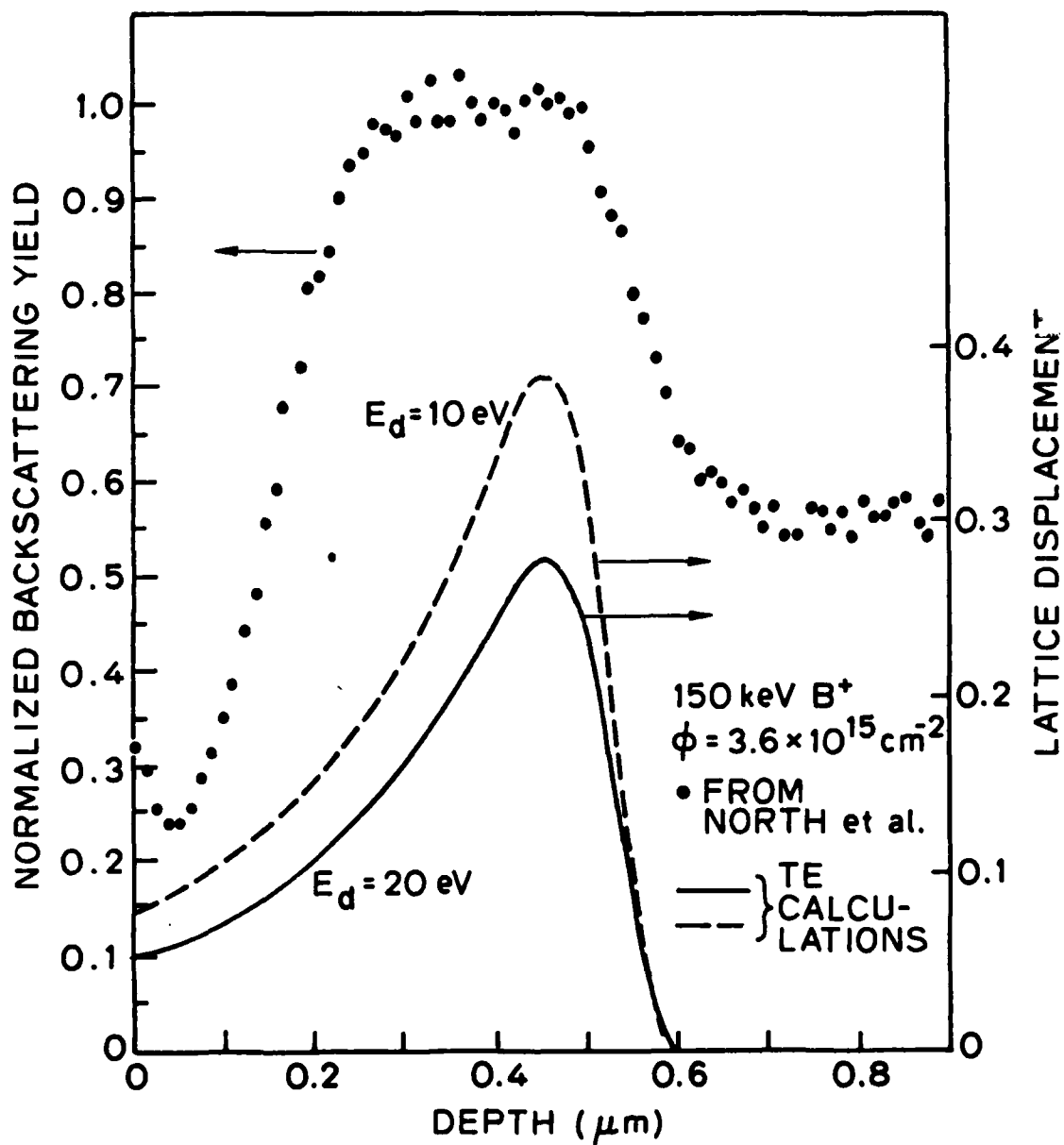
Fig. 12.3: Oxygen recoil yield as a function of oxide thickness for implantations of phosphorus into SiO_2 on silicon. Arrows indicate the projected range R_p for each energy.

Fig. 12.4: Oxygen recoil yield as a function of oxide thickness for implantations of arsenic into SiO_2 on silicon. Arrows indicate the projected range R_p for each energy.

323

Fig. 12.5: Comparison of the experimental RBS results of North and Gibson [12.20] and a TE calculation which shows the correlation between the experimentally observed edges of the buried amorphous layer and the calculated fractional displacement of the silicon lattice for an implantation of 3.6 10^{15} cm^{-2} 150 keV boron into silicon held at liquid nitrogen temperature.
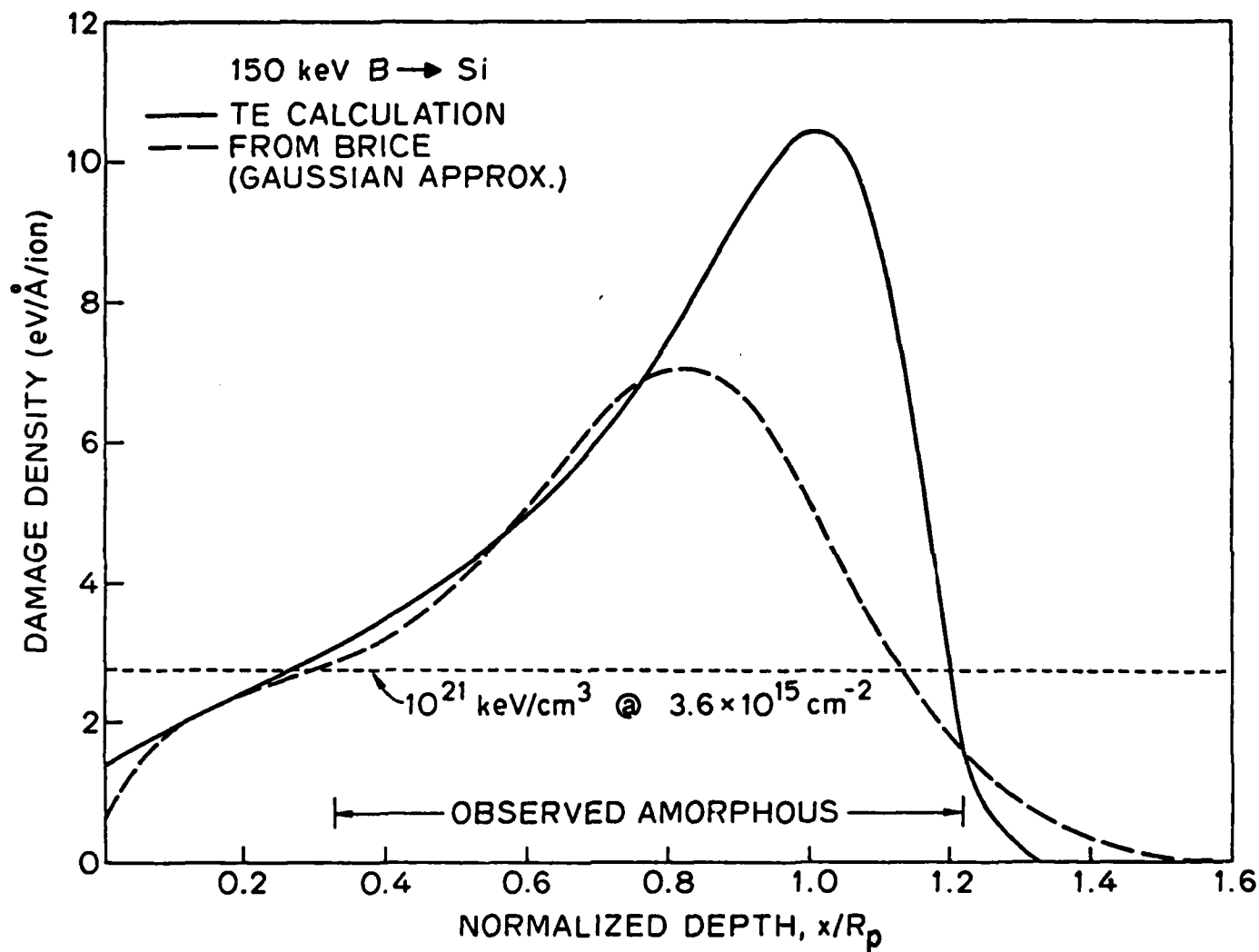
Fig. 12.6: Comparison of a TE calculation and a calculation from Brice [12.16] for the energy density deposited into elastic collision processes under the same implant conditions as in Fig. 12.5.
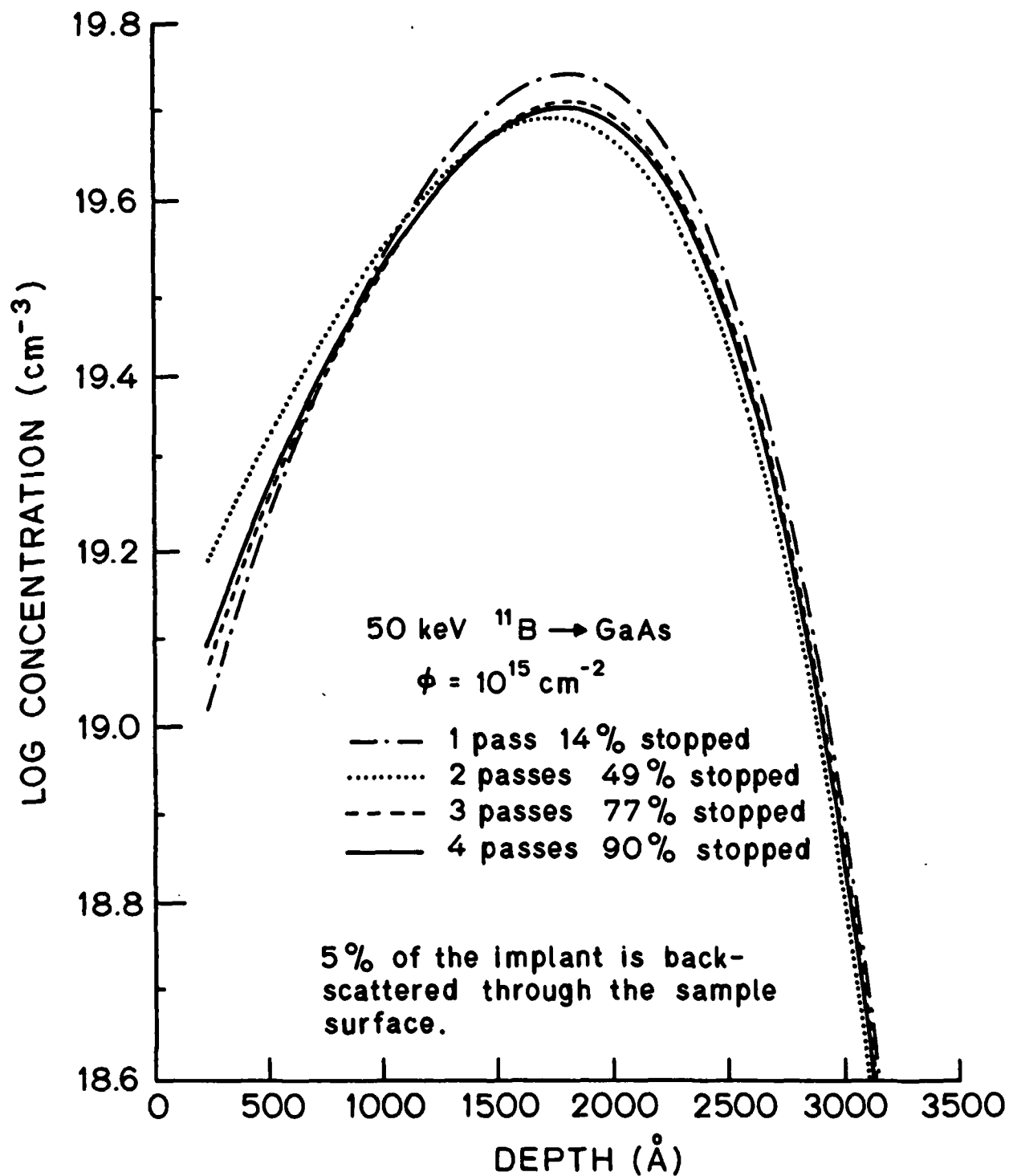
Fig. 12.7: Multiple pass TE calculation for the range distribution of 10^{15} cm^{-2} boron implanted at 50 keV into GaAs.

13. GETTERING AND TRANSIENT PROCESS CHARACTERIZATION

G. Bronner, C. Ho, J. Plummer

13.1 INTRODUCTION

The ability to efficiently follow a phenomenon's time behavior has proven a useful tool in the search for better I.C. process models. In this section we briefly review the work of previous years which showed the feasibility of using a single sample to characterize a time dependent phenomenon. We then explain in detail the work of this past year, which has concentrated on applying this technique to the problem of gettering. Finally we discuss what the results suggest for the future.

In order to study time behavior, investigators often resort to the use of multiple samples, with each sample corresponding to a different point in the time course of a transient phenomenon. This approach has a number of drawbacks.

First, multiple samples may require much processing. This can lead to use of a limited number of samples or time points. In addition, undesirable variations from sample to sample may occur due to an unknown cause. This may lead to misinterpretation of the data. Such data, when put into SUPREM, would have limited predictive power if used to simulate new processes.

Ideally, a single sample should be monitored vs. process time and temperature "in situ", as, for example, in the use of automated ellipsometry to monitor oxidation kinetics [13.1]. In electrical techniques requiring electrical contacts or interconnects, present aluminum metallization is not compatible with the high temperatures of most processes being characterized. An electrical contact technology capable of withstanding high process temperatures, and conventional chemicals used in sample preparation, would be extremely useful

327

in this work. Additionally, this contact technology should not interfere with the transient high temperature processes that will be monitored--i.e. the technology must be relatively low temperature and should not introduce electrically active impurities into the sample. Such a contact metallurgy allows a single sample to be repeatedly heat cycled with interspersed electrical measurements, as illustrated in Fig. 13.1. From such measurements a great deal can be learned about the time/temperature behavior of transient processes.

We proposed tungsten silicide (WSi_2) as a material which fulfilled the requirements. To test its feasibility, wafers were processed incorporating MOS capacitors and n^+-p diodes with WSi_2 metal lines. These wafers were then annealed at temperatures ranging form 700°C to 1100°C. The capacitors were measured to see how WSi_2 affected the SiO_2/Si interface and thus capacitor quality. The results were promising. The n^+-p diodes were measured for leakage currents to check whether or not impurities were diffusing from the silicide into the bulk Si. WSi_2 contacts did not seem to affect the n^+-p diodes in any way. When used as a gate electrode WSi_2 also performed well. The sputtering process by which the silicides were deposited did seem to introduce some damage to the SiO_2/Si interface. However, the damage annealed out at high temperature leaving the WSi_2 gate capacitors indistinguishable from Al gate devices.

Confirming these results is another experiment. Bipolar transistors were fabricated using WSi_2 metal lines. These devices should be more sensitive to metallic contamination than p-n diodes. In a p-n diode one has diffusion currents which can increase the leakage currents seen due to effects far away from the device. In the bipolar device, any excess base currents can be attributed solely to contamination in the base-emitter space charge region.

The bipolar devices were fabricated and immediately after metal deposition the solid curve in Fig. 13.2 was measured. The wafer was then cleaned and

annealed at 850°C for 1 hour. The same device was again measured. It is also plotted in the figure, this time as a dot/dash line. We see that the collector currents are unchanged while the base current is increased at low levels. We believe this rise is due to trace contamination from the metal tweezers used in handling the wafer. Indeed as we measured devices away from the periphery of the wafer we found no change in device characteristics. This is important in several ways. It establishes our ability to detect low levels of contamination. Additionally, it demonstrates our ability to study the performance of the same device after annealing.

13.2 GETTERING AS A TRANSIENT PHENOMENON

During the past year, work has concentrated on applying this technique to the problem of backside damage gettering. Since gettering is believed to involve silicon point defects, it was thought that characterization of this process would provide useful insight into the role of these defects in an interaction of damaged layers with other processes such as diffusion and oxidation.

To get reasonable yields from large circuits, processing must be as clean as possible. Studies have shown that even in the cleanest laboratories, contamination can be introduced from sources such as tweezers, quartzware, and furnace tubes [13.2]. A method frequently used to remove unwanted contaminants is backside damage gettering. The aim of gettering is to draw impurities away from active device areas (where they degrade circuit performance) to nonactive areas such as the backside of a wafer. See Fig. 13.3. Although gettering has been studied extensively [13.3-13.6], it is still not well understood.

To monitor the gettering action with time we chose the bipolar transistor. Looking at Fig. 13.3, the region where metallic impurities are least desireable is the base-emitter space charge region. Impurities here give rise to generation recombination currents [13.7], which degrade performance at low current levels.

329

This effect can be seen in Fig. 13.4. Simple theory predicts that both the collector current, I_c, and the base current I_b, should be exponentially dependent upon the base-emitter voltage, V_{be} such that I_c or $I_b \propto \exp$ $\{qV_{be}/kT\}$. In fact, the collector current, I_c, does behave that way, but the base current does not. One finds the following dependence: $I_b \propto \exp$ $\{qV_{be}/mkT\}$. As one scans across a wafer measuring many devices, one finds a distribution in the base currents. As predicted by the Shockley, Read, Hall recombination theory [13.8], one finds the m-factors in the exponential varies such that $1 < m < 2$. When $m = 1$ we have negligible contamination, while the case of $m = 2$ corresponds to contamination by a metal which introduces a trap level into the middle of the silicon bandgap. Thus, by looking at the base currents across the wafer we hope to track the contamination level as a function of time.

Qualitatively this can be done for the wafer characterized in Fig. 13.4. However, quantifying these results, i.e. correlating a certain base current specifically to some N cm^{-3} metal atoms, is difficult. For the wafer shown in Fig. 13.4 this is not possible due to the unknown nature of the contaminants. If, on the other hand, a sample is intentionally contaminated with a well behaved metal, the task may be more amenable. Gold is nearly ideal for these purposes. It introduces a trap level very close to the middle of the silicon bandgap [13.9]. Additionally, there is a large body of work on gold diffusion in silicon [13.10-13.12]. This information base makes interpretation of results possible.

For our first gettering experiment, a wafer was processed with bipolar transistors on it. As outlined in the start of this report, WSi_2 was used as the metal layer. After normal fabrication was completed a thin layer of gold was evaporated on the back of the wafer and then driven in at 850°C.

Once the wafer was gold doped, a backside damage region was created using a Nd-YAG laser on half of the wafer. Recent studies have shown laser damage to be an effective getter which is stable through high temperature processing [13.13]. The wafer was measured before and after anneals of 1/2 hour, another 1/2 hour, 1 hour, and 4 hours. All the anneals were done in nitrogen at a temperature of 850°C. In Fig. 13.5 we have plotted normalized N_t (the level of gold contamination) as a function of anneal time. We see that N_t drops sharply after a 1/2 hour heat treatment and then gradually relaxes back to its original value after long anneal times.

This behavior is contrary to expectations. Once gettering action has begun, one expects device performance to improve monotonically until all impurities have been removed from the device. Possible explanations of the data are: 1) The backside damage is being annealed out at 850°C and the gold, after initially being trapped at the backside, is being released. This seems unlikely due to the reported stability of laser damage. 2) Gold is being gettered to the backside. However, additional impurities from the furnace and/or the WSi_2 layer are degrading the devices. 3) Gettering action is not occuring. Instead the change in device characteristics is just the gold reequiliberation.

Of these possibilities we feel the last is most plausible. When gold diffuses into a wafer, it moves primarily as an interstitial atom, since interstitial diffusion is much faster than substitutional diffusion. However, once equilibrium is reached, gold prefers to sit substitutionally on the silicon lattice. This requires the reaction of an interstitial gold atom with a silicon vacancy such that: $A_{U_I} + V \rightarrow A_{U_{SUB}}$. To the gold atoms, the free silicon surfaces look like sources for vacancies. This leads to the pile-up of gold at the surfaces which is commonly referred to as a "U-shaped" profile (see Fig. 13.6). To prevent this pile-up from saturating the backside damage region with

331

gold, approximately 10 microns of the back surface was etched away prior to the laser damage. These factors may explain the experimental results above. During the first anneal, gold is driven towards the back side as a surface pile-up is re-established. This results in a depletion of gold from the front of the wafer. During subsequent anneals, as equilibrium is approached, gold returns to the front of the wafer with concentrations similar to those seen initially.

Thus, from this experiment, conclusions are difficult to draw. Additional information might be obtained by analysis of gettering in an uncontaminated sample. By changing the times and/or temperatures of the anneals it is hoped that we can determine what conditions are necessary for effective gettering. In addition, these tests should tell us whether the wafer was being contaminated by the furnace and/or the silicide layer.

A second experiment was pursued to look at these questions. As in the previous experiment, a wafer was processed with bipolar transistors on it, with WSi_2 used for metallization. After the normal fabrication was completed, half of the wafer was backside damaged using a Nd-YAG laser. The wafer was then annealed in a nitrogen ambient at 850°C for times of 1/2 hour, 1/2 hour, 1 hour, 2 hours, 4 hours, and 8 hours. After each anneal the wafer was rapidly cooled to room temperature, given a 1/2 hour forming gas anneal at 450°C, and then electrically characterized.

To evaluate the effectiveness of the gettering, a number of identical devices across the wafer were characterized. As previously discussed, the base current of any device at low current levels can be modelled such that $I_b \propto \exp\{qV_{be}/mkT\}$. Due to the statistical distribution of the impurities, the base currents show a spread which can be modelled by a distribution in m-factors. If gettering is taking place, we would expect the m-factors to approach 1 and the spread in the base currents to disappear as the impurities

332

are drawn from the base-emitter space charge region. In Fig. 13.7 we have plotted the results. Initially the devices show m-factors in the range of 1.13 to 1.56 with a mean of 1.3. Subsequent annealing increased the mean to 1.6 while decreasing the spread. From this we can conclude several things. First, it is obvious that the furnace and the WSi_2 layer are not contaminating the devices in any significant way. It is also clear that for the times and temperature used in this experiment (850°C for up to 16 hours) gettering was not yet evident. And our ability to follow the time behavior of our devices was again demonstrated.

Since gettering did not seem to be taking place at 850°C, it was decided to repeat the same experiment, this time performing the anneals at 950°C. As in the previous case, devices were fabricated and then backside damaged with a Nd-YAG laser. The wafer was then annealed for times of 1/4 hour, another 1/4 hour, 1/2 hour, 1 hour, 2 hours, and 4 hours. Again the data from this experiment was reduced to the form of a plot of m-factors vs. anneal time (see Fig. 13.8). Initially the characteristics are scattered with m-factors between 1.5 and 1.7 with a mean of 1.6. With subsequent anneals the mean finally rises to 1.73 with the scatter in values decreased considerably.

Thus in these two experiments we have been unable to see any gettering of metallic impurities at either 850°C or 950°C for times up to 16 hours. There are several possible explanations for this:

1) Perhaps gettering is simply not occurring in the time/temperature range we are using. This is a possibility since we used laser backside damage as our gettering agent. This is a new process which has not been well characterized. We may simply be using inappropriate parameters. To avoid this uncertaintly, the next set of experiments are based on an Argon Ion Implantation damage gettering scheme which is known to getter effectively.

2) Another possibility is that gettering is taking place but competing contamination is masking its effect. This seems unlikely considering the stability of the device characteristics during long anneals in the previous experiments. It seems unlikely that a quasi-equilibrium situation would be reached whereby the amount of metallic impurities drawn to the backside getter is balanced by the amount of new contamination entering the device. Nevertheless, to check this possibility, the next set of devices fabricated will have both bipolar and MOS devices. The MOS devices with their protective gate oxides, should be immune from contamination. Thus by comparing MOS and bipolar characteristics we can determine if contamination is a problem.

3) Finally, we must ask if we are really looking at metallic contamination when we examine the behavior of the base current in bipolar transistors. Vacancies are hypothesized to affect many processes in silicon. Indeed the models used in SUPREM to explain oxidation of heavily doped silicon and the diffusion of heavily doped regions depend on these formalisms. Recently, it has been suggested that vacancies can affect device performance [13.14]. This seems plausible. Vacancies are known to introduce levels into the silicon bandgap. See Fig. 13.9. At 950°C the total number of vacancies in lightly doped silicon V_T is 2.7 x 10^{12} cm^{-3}. At 850°C, V_T is 8.6 x 10^{11} cm^{-3}. In doped regions these values are even higher. If these vacancy levels were frozen in and if they were electrically active, they could dominate the generation-recombination processes and swamp out any effects due to metallic contamination.

To test this idea, we performed the following experiment. A wafer was oxidized at 1000°C to passivate the surface. It was then annealed at temperatures ranging from 600°C to 1000°C. After each anneal the wafer was cooled to room temperature as quickly as possible. It was then given a 1/2 hour anneal

in forming gas at 450°C. Bulk lifetime was then measured by looking at the change in conductivity as a light was strobed on and off.

If frozen-in vacancies actually limit device performance, we should see the lifetime dropping off with higher anneal temperatures. A plot of lifetime vs. inverse temperature (τ vs. 1/T) should extract the activation energy of vacancy formation. This is done in Fig. 10. We see that from 700°C to 1000°C the lifetime is essentially constant. Only at the lowest anneal temperature, 600°C, is the lifetime somewhat higher. From this we conclude that frozen-in vacancies do not seem to affect device behavior. Thus, when we look at low current operation of bipolar transistors we are indeed looking at a phenomenon caused by metallic impurities.

We should qualify the above statement, noting that the result is obtained by looking at recombination in the bulk and it then is being applied to the case of generation in a depletion region. This is valid if Shockley, Read, Hall recombination is valid. To check this, the same experiment is to be repeated using lifetime in MOS capacitors, which will test lifetime in a depleted region.

Beyond that, we have a number of plans for the future. Our first is to process a wafer fully, including a gettering technique we know to work well. Another wafer will receive the same process, except for the gettering treatment. It will receive the same backside damage, but instead of one long anneal, it will receive several shorter ones with its device characteristics being monitored along the way. Thus we will see if five one-hour anneals, for example, is the same as one five-hour anneal.

If the results of the above experiment are positive, we will then intentionally contaminate another wafer with gold. We will then try to getter the gold, tracking this gettering through its electrical effects, and thus

quantifying it. We will then attempt to characterize gettering using various backside treatments. People have reported using many forms of damage to getter impurities. Whether gettering using these is a fundamentally different process, or if they all represent minor variations on the same theme, is a major project we hope to undertake.

We would like to point out that these studies should provide useful information on the ways in which damaged layers interact with various silicon point defects and other processing steps. The comparison between phosphorus diffusion gettering and the other types of damage gettering should be particularly fruitful in this regard. This stems from the work of various persons studying heavy doping effects in phosphorus diffusion. These studies have led to an understanding of the interactions between phosphorus and silicon vacancies, which are believed to be the dominant defect in heavily doped phosphorus layers. Similarities and differences in gettering behavior should help establish the ways in which damaged layers interact with point defects.

13.3 SUMMARY

We have shown that the use of WSi_2 metallurgy allows us to characterize efficiently a high temperature transient process. Therefore, we have begun to use the technology to address the many problems included under the heading of gettering. While our initial experiments have been inconclusive, they have still taught us a great deal. As this knowledge is applied in the next generation of experiments, we believe that much data and insight related to the interactions of heavily damaged regions and silicon point defects will be gained. This understanding should be crucial to understanding the interdependencies of process steps, and lead to much improved models in SUPREM III. All of the diffusion and oxidation models used in SUPREM III depend upon knowledge of local and instantaneous concentrations of interstitials and vacancies. The

336

techniques used in this project should provide vital information about the
behavior of these defects in silicon.

REFERENCES

[13.1] See the section of this report on Thin Oxide Kinetics.

[13.2] P. F. Schmidt and C. W. Pearce, "A Neutron Activation Analysis Study of the Sources of Transition Group Metal Contamination in the Silicon Device Manufacturing Process," J. Electrochem. Soc., 128, 630 (1981).

[13.3] T. E. Seidel, R. L. Meek, and A. G. Cullis, "Direct Comparison of Ion-Damaged Gettering and Phosphorus-Diffusion Gettering of Au in Si," J. of Appl. Phys., 46, 600 (1975).

[13.4] R. L. Meek, T. E. Seidel, and A. G. Cullis, "Diffusion Gettering of Au and Cu in Silicon," J. Electrochem. Soc., 122, 786 (1975).

[13.5] L. Baldi, G. Cerofolini, and G. Ferla, "Heavy Metal Gettering in Silicon-Device Processing," J. Electrochem. Soc., 127, 164 (1980).

[13.6] M. J. T. Lo, J. G. Skalnik and P. F. Ordung, "Implantation Gettering of Gold in Silicon," J. Electrochem. Soc., 128, 1569 (1981).

[13.7] C. T. Sah, R. N. Noyce, and W. Shockley, "Carrier Generation and Recombination in p-n Junction and p-n Junction Characteristics," Proc. IRE, 45, 1128C (1957).

[13.8] W. Shockley and W. T. Read, Jr., "Statistics of Recombination of Holes and Electrons," Phys. Rev., 87, Sept. 1952, pp. 835-842.

[13.9] F. Richou, G. Pelous, and D. Lecrosnier, "Thermal Generation of Carriers in Gold-Doped Silicon," J. Appl. Phys., 51, 6252 (1981).

[13.10] W. M. Bullis, "Properties of Gold in Silicon," Solid State Electronics, 2, 143 (1966).

[13.11] F. A. Huntley and A. F. W. Willoughby, "The Effect of Dislocation Density on the Diffusion of Gold in Thin Silicon Slices," Solid State Electronics, 13, 1231 (1970).

[13.12] F. A. Huntley and A. F. W. Willoughby, "The Effect of Dislocation Density on the Diffusion of Gold in Thin Silicon Slices," J. Electrochem. Soc., 120, 414 (1973).

[13.13] Y. Hayafuji, T. Yanada, and Y. Aoki, "Laser Damage Gettering and Its Application to Lifetime Improvement in Silicon," J. Electrochem. Soc., 128, 1975 (1981).

[13.14] J. G. Fossum and D. S. Lee, "Significance of Low-Temperature Processing in the Fabrication of Silicon Solar Cells," Record of Fifteenth IEEE Photovoltaic Specialists Conf. - 1981, [Cat. No. 81CH1644-4], p. 120.

[13.15] R. B. Fair and J. C. Tsai, "A Quantitative Model for the Diffusion of Phosphorus in Silicon and the Emitter Dip Effect," J. Electrochemical Soc., 124, 1107 (1977).
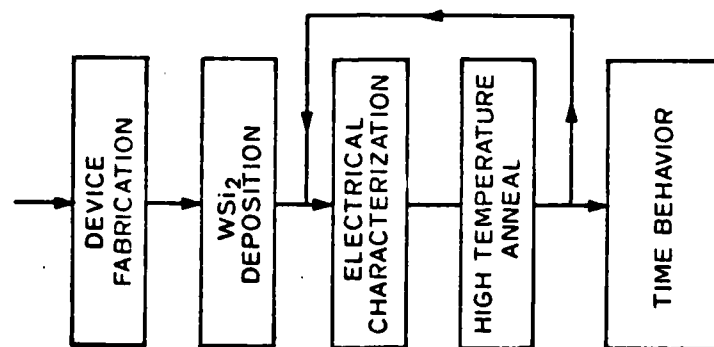
Fig. 13.2: Low base and collector current behavior of bipolar test devices showing effect of metallic contaminants on base current voltage dependence.

Fig. 13.1: Schematic representation of experimental procedure aimed at studying transient phenomena.
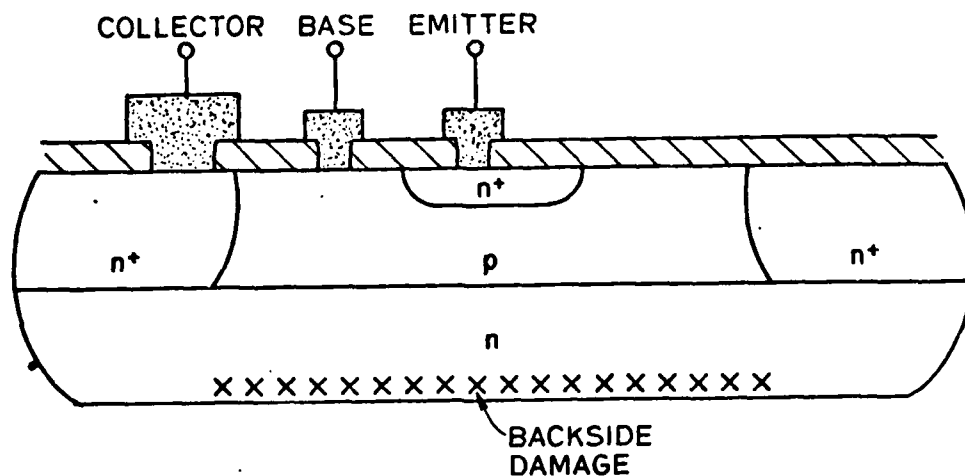
Fig. 13.3: Bipolar transistor structure used to study gettering. Gettering involves the movement of metallic impurities away from the active device areas of the wafer.
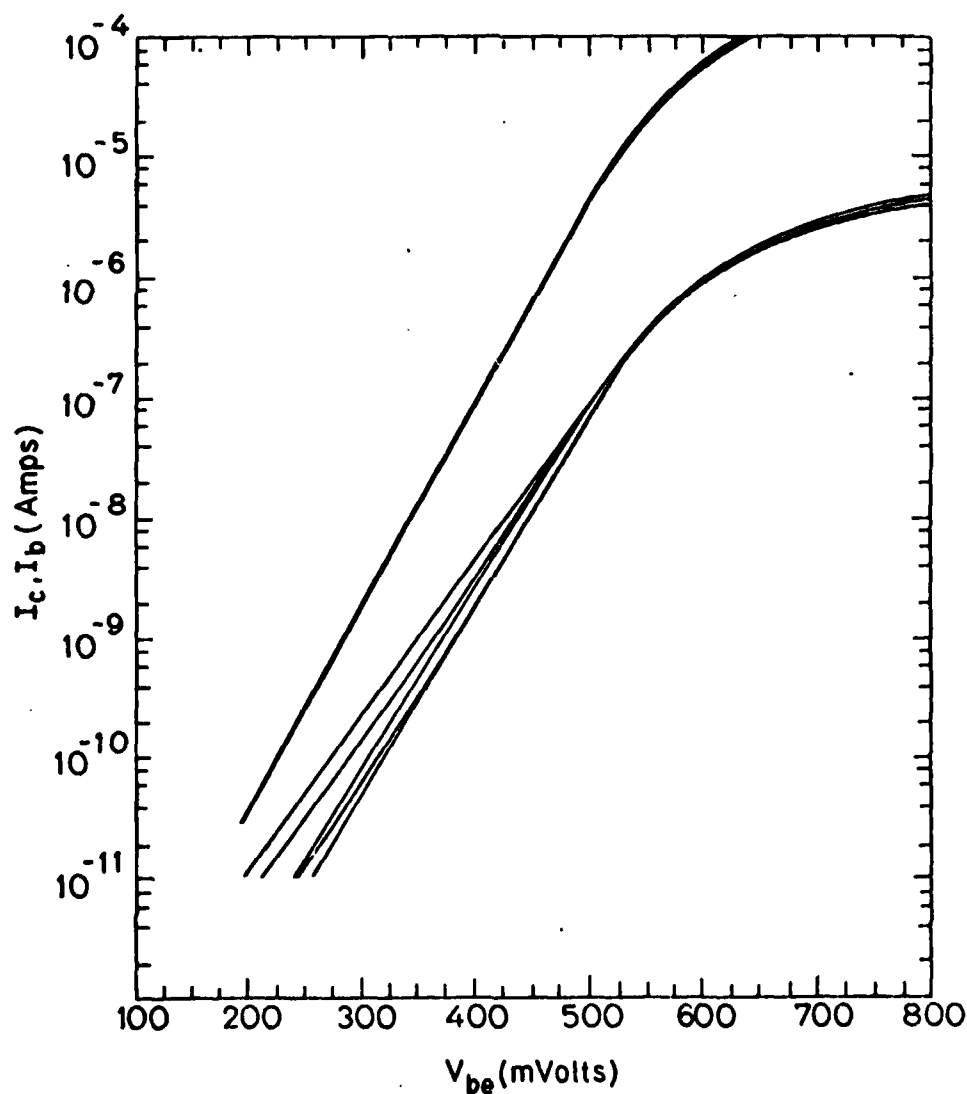
Fig. 13.4: I_c and I_b for a number of NPN transistors across a wafer.
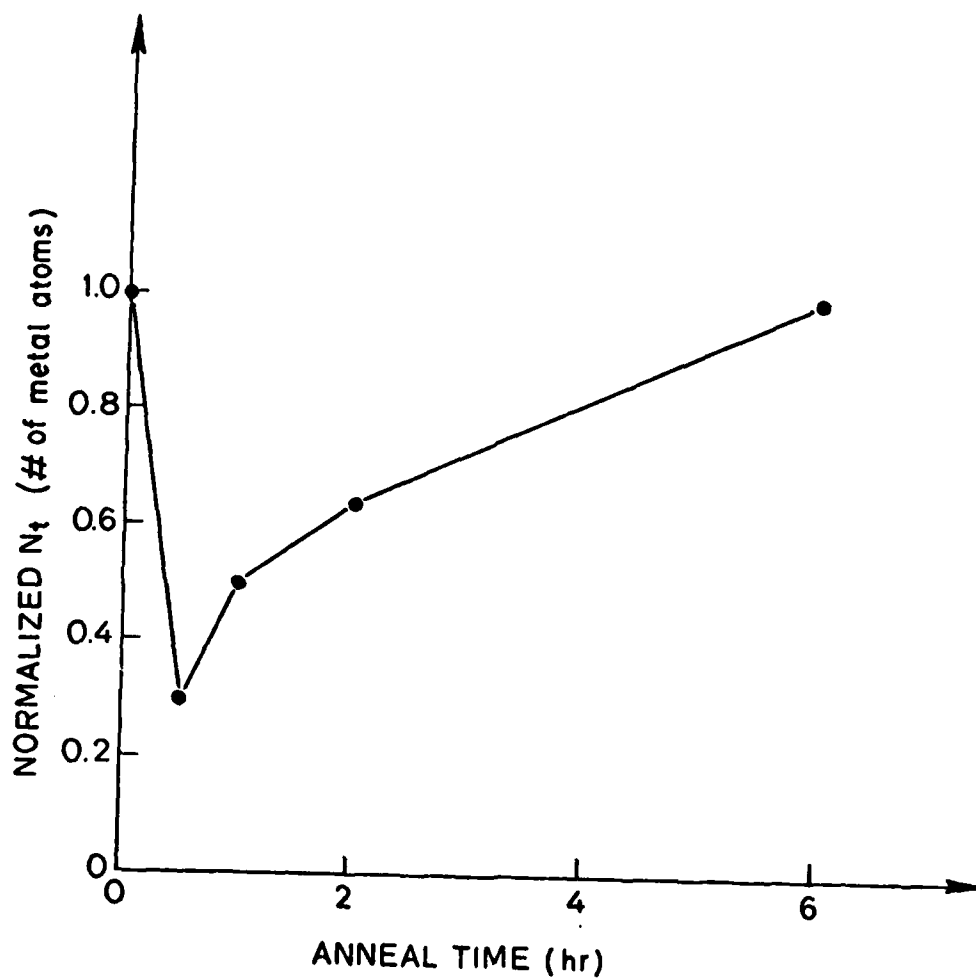
Fig. 13.5: Normalized density of metallic impurities N_t as a function of gettering time.
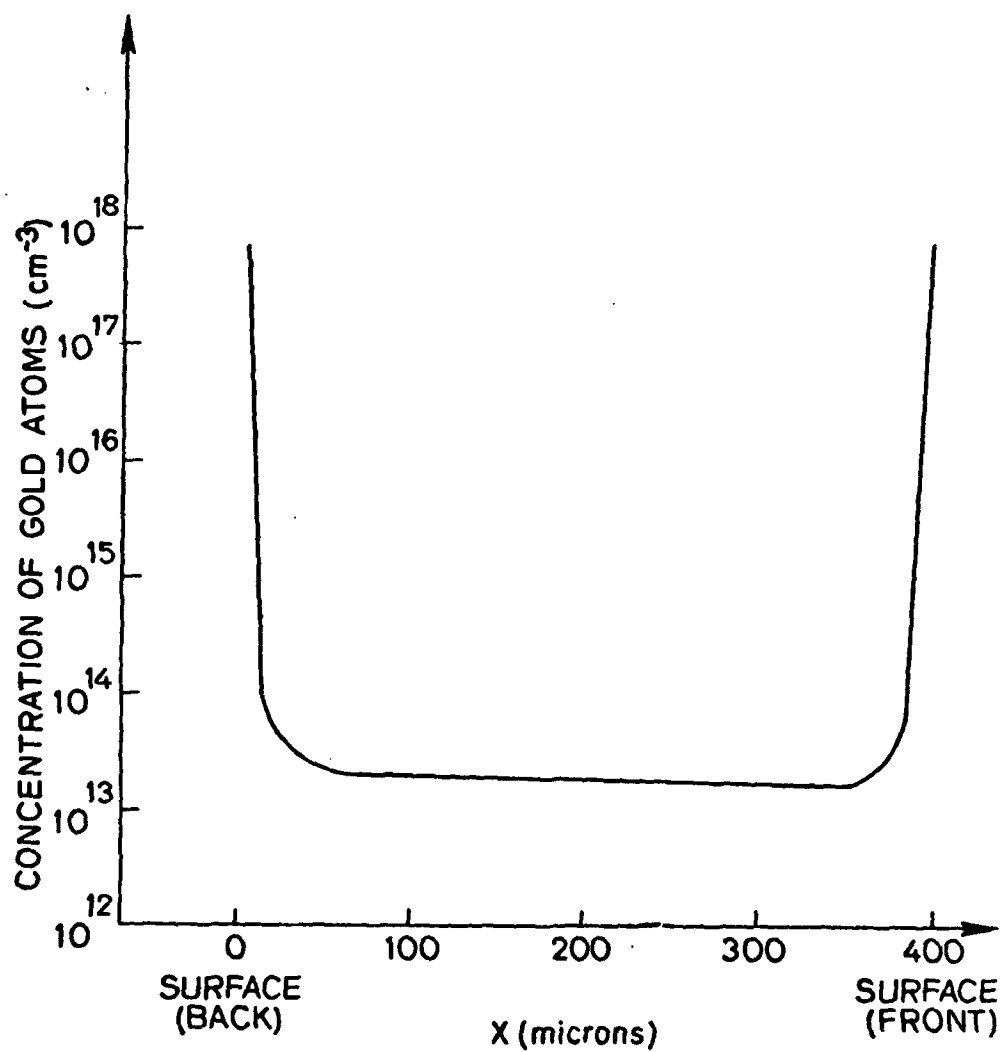
Fig. 13.6: Concentraion of Gold as a function of depth into the wafer.
Temperature of drive-in assumed to be 850°C.

342
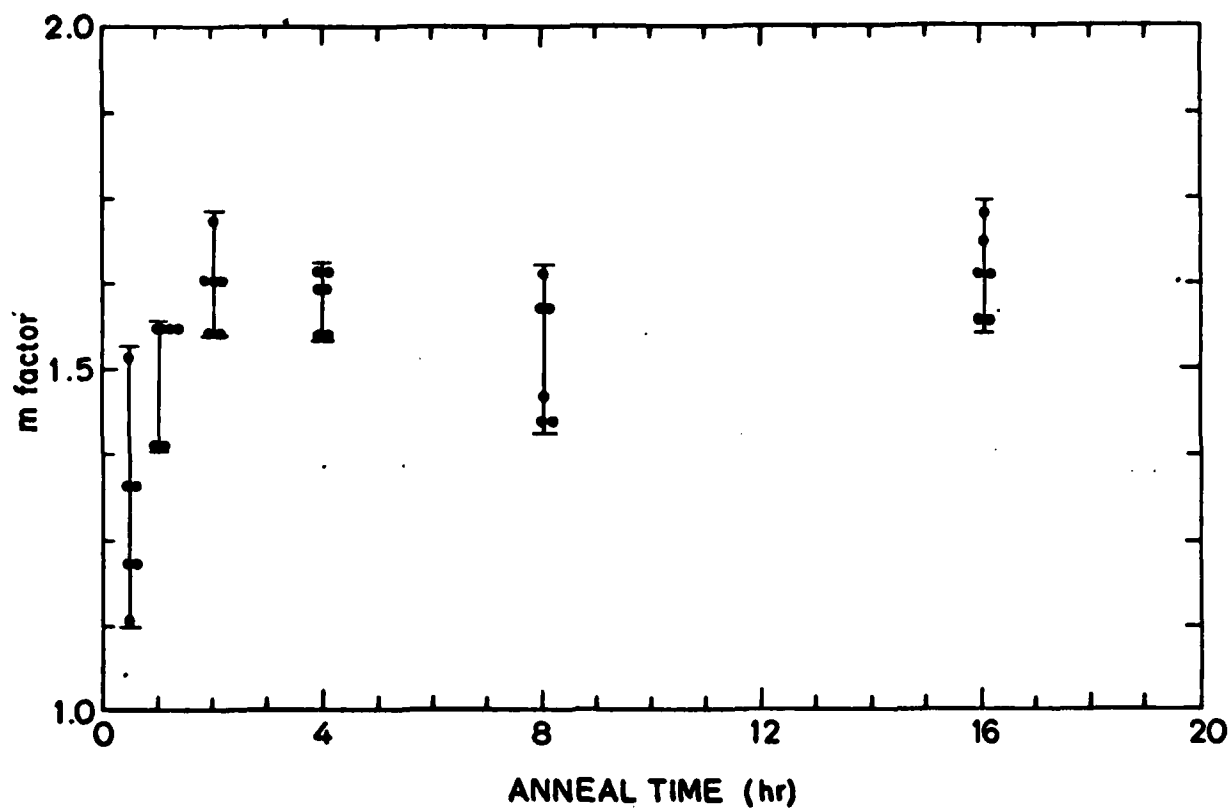
Fig. 13.7: Low current I_b slopes (m) of devices after successive gettering anneals at 850°C.
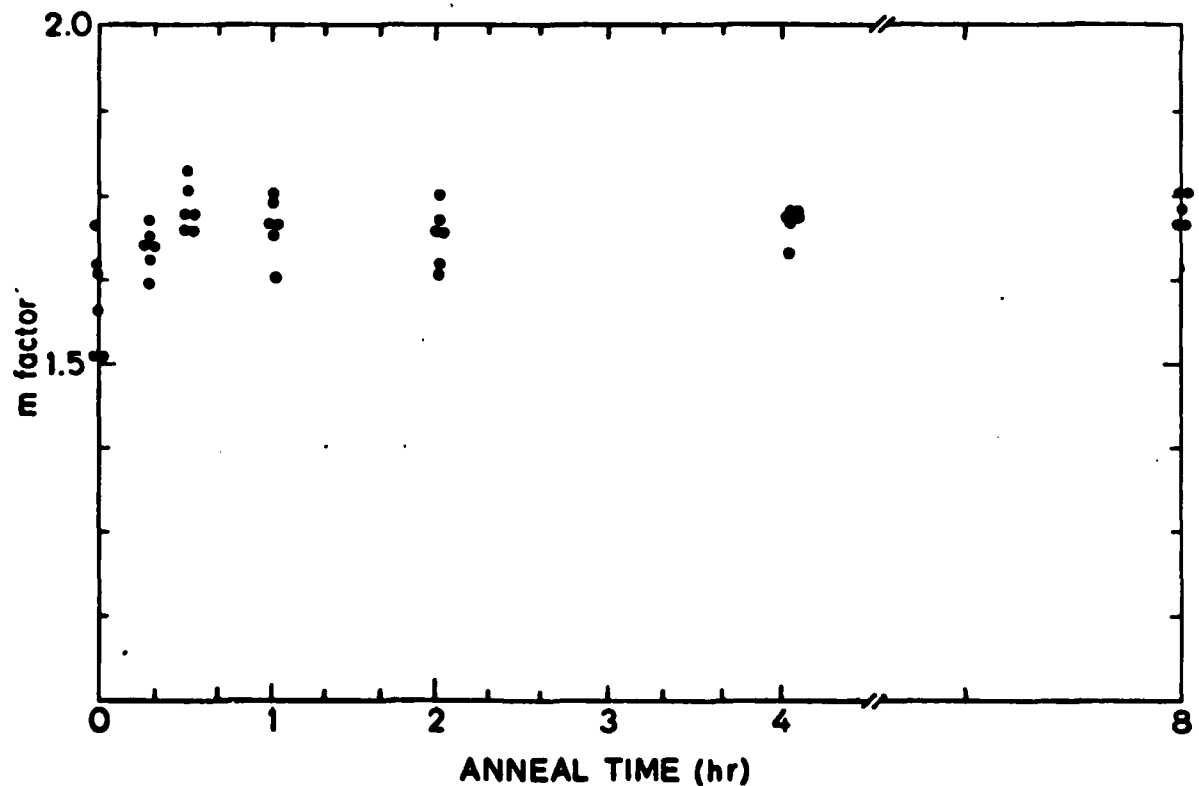
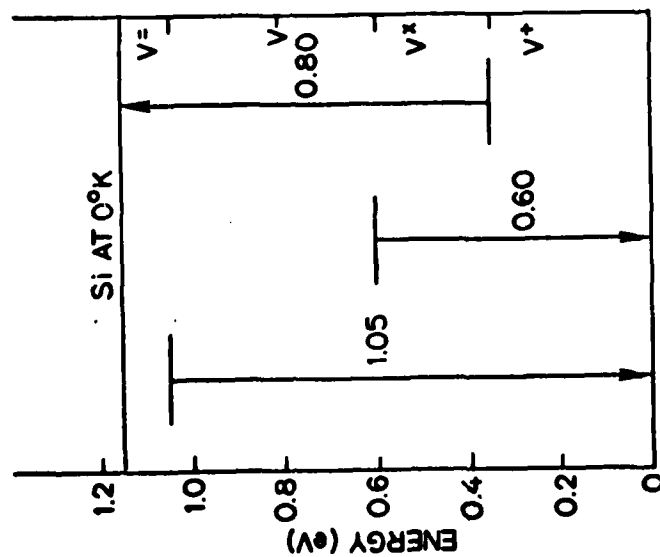Fig. 13.8: Low current I_b slopes (m) of devices after gettering anneals at 950°C.

343

Fig. 13.10: Bulk Lifetime, τ, measured after quenching sample from temperature T.

Fig. 13.9: Vacancy energy levels in silicon at 0°C.

344

14. POLYSILICON AND SILICIDE CONTACTS TO SILICON

B. Swaminathan, F. Shone, K. C. Saraswat, R. W. Dutton, J. D. Plummer

14.1 INTRODUCTION

Ohmic contacts to silicon in a device structure have traditionally been made using metals such as Al, Ti, Mo, etc. In this type of technology, contact formation is done towards the end of the device fabrication. All of the high temperature steps are completed before the metal to silicon contacts are formed. After the contacts are formed only low temperature fabrication steps are allowed because the metal to semiconductor contacts cannot withstand high temperatures. During the last few years several new technologies have been emerging which utilize entirely different techniques to form contacts. In these technologies the contacts are formed early in the process sequence and thus are subjected to several high temperature steps. The motivation for this change has generally been denser device structures. This change modifies the requirements imposed on the materials being used to form contacts. Because of increasing use and interest in these types of contacts, it is important to develop suitable process models for SUPREM.

Doped polysilicon is increasingly being used to form contacts to source and drain junctions in MOS transistors [14.1] and to emitters in bipolar transistors [14.2]. In these applications the doped polysilicon serves two roles. First, it acts as a source of dopant to form the junction and then it also forms the contact to the junction. As a result the polysilicon contact is self-aligned to the junction. No additional contact hole opening is required in this scheme. Thus the junction area is minimized resulting in an overall reduction in device size. Contact to the polysilicon can be made with conventional metals like Al or by refractory metal silicide layers. This contact is generally made in an area of

the circuit where the material underneath the polysilicon is thick SiO_2. In other variations of this scheme the junction is first formed by conventional techniques and a silicide or polycide (silicide/polysilicon combination) layer is simply used to form the contact to it. The use of doped silicide contacts directly to silicon has also been suggested [14.3,14.4] where the doped silicide acts as a source of dopant diffusion to form the junction and simultaneously forms an ohmic contact as well. This technique is similar to the use of doped polysilicon, however, it has the advantage of lower silicide resistivity for interconnection applications.

In these and other similar situations the silicide/silicon and silicide/ polysilicon/silicon interfaces are subjected to high temperatures after their formation. This allows the dopant to move in and out of the various layers and to redistribute accordingly. There are very serious implications of this behavior. It has been observed that if $MoSi_2$ contacts to N^+ shallow junctions are annealed above 700°C the contact resistance increases very rapidly [14.3]. The higher the anneal temperature, the higher the contact resistance becomes. If instead the silicide is doped heavily with phosphorus the contact resistance remains low [14.4]. Tsai et. al. [14.5] and Pan et. al. [14.6] have observed that if WSi_2 contacts to doped thin polysilicon films are subjected to high temperatures, a substantial amount of dopant can diffuse from the polysilicon to the WSi_2. At the WSi_2/polysilicon interface the concentration is not uniform. There are probably both segregation and redistribution effects taking place. Again there are important device implications of these effects.

In this report we described the details of dopant diffusion from doped polysilicon into single crystal silicon to form polysilicon contacts to silicon, and the effects of oxidizing ambients during diffusion. Dopant segregation at the and poly-Si/Si interface have been taken into account. A brief description of silicide to silicon contacts is also given.

14.2 POLYCRYSTALLINE CONTACTS TO SINGLE-CRYSTAL SILICON

Polysilicon is now finding increasing use as a source for diffusing dopants into single-crystal silicon both in bipolar and MOS circuits. Novel devices have been proposed in which the self-alignment features made possible by the use of polysilicon contacts have been exploited. This report discusses the work done in 1981-1982 on characterizing dopant diffusion from a polysilicon doping source to single-crystal silicon. This is a continuation of work reported in the 1981 annual report [14.7].

14.2.1 Polysilicon Doping Source in an Oxidizing Ambient

It was shown in the 1981 annual report [14.7] that in the diffusion of dopant from a polysilicon doping source, an important effect is the absorption of interstitials in the grain boundaries in the polysilicon. The silicon self-interstitials are generated at the oxidizing interface, and any which reach the substrate enhance the diffusivity of dopants there. Thus the oxidation enhanced diffusion is strongly dependent on the structure and thickness of the polysilicon. The concentration of interstitials in the substrate has been postulated to be

$$(C_I - C_I^*)_{substrate} = (C_I - C_I^*)_{surface} \ exp\left(\frac{-Z_{poly}}{L_I}\right) \qquad (14.1)$$

Here C_I is the concentration of interstitails induced by the oxidation and C_I^* is the equilibrium intersitial concentration. Z_{poly} is the thickness of polysilicon still unoxidized and L_I is a characteristic adsorption length for interstitials in polysilicon. From this the oxidation enhanced diffusivity, D may be written as follows.

$$D = D_o + A_I K\left(\frac{dx}{dt}\right)^{0.4} exp\left(\frac{-Z_{poly}}{L_I}\right) \qquad (14.2)$$

347

In this expression D_0 is the diffusivity in the absence of oxidation and includes high concentration effects, A_IK is an empirical constant and dx/dt is the oxidation rate.

A. Confirmation of Interstitial Absorption

The experiments described in the 1981 report [14.7] did indeed show the absorption of silicon self-interstitials at the grain boundaries. But these experiments involved seperating the effects of high-concentration diffusion and movingboundary oxidation from that of oxidation enhanced diffusion. An additional experiment was therefore performed to demonstrate in a simple and incontrovertible manner the interaction of interstitials with grain boundaries. In this experiment the polysilicon layer was not doped and, instead, the diffusion in a phosphorus buried layer was studied. The starting wafers were predeposited with phosphorus from a $POCl_3$ source at 950°C for 30 minutes. The phosphorus was then driven-in by a heat treatment at 1200°C for 30 minutes in oxygen. A 5 μm thick boron-doped epitaxial layer was grown on the samples in an atmospheric pressure reactor using silane at 1050°C. Polysilicon was then deposited on the wafers by reducing the temperature of the reactor to 750°C. Two thicknesses of polysilicon were deposited - 3000 Å and 6000 Å. Next, successive layers of oxide, nitride and oxide were deposited in the APCVD reactor. Stripes of 1 mm wide nitride regions alternating with 1 mm wide bare regions were obtained by photolithography. The samples were subjected to oxidation for various times in oxygen at 1000°C. The differential diffusion of the buried layers under oxidizing and non-oxidizing conditions was revealed using spreading resistance measurements and by using a copper sulfate solution which stains the n-type region. The resulting structure is shown in Fig. 14.1 and it is expected that the regions that are oxidized will have diffused more than the regions under nitride as schematically illustrated in the figure.

348

Figures 14.2(a) and 14.2(b) show spreading resistance measurements of concentration profiles in some of these buried layer structures which had been oxidized in dry O_2 for 17.5 hours. The greater diffusion in the region and that was oxidized is clearly evident in Fig. 14.2(a) for the sample with 3000 Å of polysilicon. In contrast in Fig. 14.2(b) which is for the sample with 6000 Å of polysilicon, the concentration profiles in both the oxidized and unoxidized regions are nearly the same and are also very similar to the profile in the unoxidized region of the sample with 3000 Å of polysilicon. The difference between the two samples is clearly caused by the absorption of interstitials at the polysilicon grain boundaries. The 6000 Å polysilicon film absorbs most of the interstitials that are generated by the oxidation, while the 3000 Å film allows a larger fraction of the interstitials to reach the substrate. Further confirmation was obtained by subjecting the samples to more oxidation. This reduced the thickness of polysilicon in the oxidized regions sufficiently to allow enough interstitials to reach the substrate to cause a significant difference in the diffusion in the oxidized and unoxidized regions, even for the sample with 6000 Å of oxide.

B.   Effect of Grain Size

The experiments described above have established that the grain boundaries of polysilicon are very effective in absorbing interstitials that are generated by oxidation. The area of grain boundary per unit volume in a layer of polysilicon is inversely proportional to the average grain size of the polysilicon. It may therefore be expected that smaller grained polysilicon will be more effective in absorbing interstitials than larger grained material.

The effect of the grain structure of the polysilicon doping source was studied by fabricating structures as in Fig. 14.3. The polysilicon layers were 3000 Å thick and deposited at 750°C and 900°C. Since at 900°C, silicon deposited directly on single-crystal silicon tends to form poor epitaxial layers rather than

349

polycrystalline layers, the 900°C films were deposited on 500 A nucleating layers of polysilicon deposited at 750°C. The grain sizes for the films deposited at 750°C and 950°C are estimated to be 500 A and 1000 A respectively [14.8]. The samples were implanted with 3.3×10^{14} atoms/cm^2 of phosphorus at 60 KeV and then then diffused for various times at 1000°C in wet oxygen. Junction depths were measured by angle lapping and staining.

Fig. 14.4 shows junction depths as a function of the square root of time for the two different grain-size samples oxidized in wet oxygen at 1000°C. It is seen that in the regions protected by the nitride, the junction depth increases linearly with the square root of time and is practically the same for both sizes of polysilicon grains. In the oxidized regions, the junction depth is not a linear function of the square root of time because the OED effect increases with time as the polysilicon is consumed. Moreover the junction is greater for the larger grained material as expected. Absorption lengths of interstitials of 500 A and 1000 A were extracted for the two samples.

14.2.2 Arsenic Diffusion from a Polysilicon Doping Source

Concentration profiles after arsenic diffusion from a polysilicon doping source have been studied using Rutherford backscattering spectrometry. The concentration profile has been found to be mainly determined by the diffusion in the single-crystal silicon since the diffusion in the polysilicon is so rapid that the concentration profile in the polysilicon is relatively flat. Segregation to the grain boundaries of arsenic has been found to be an important factor which causes two effects - a pile-up of arsenic at the polysilicon/single-crystal silicon interface and a difference between the surface concentration in the single-crystal silicon and the average concentration in the polysilicon. The effect of interface oxide layers between the polysilicon and the single-crystal silicon has been studied.

350

<100> oriented boron-doped silicon wafers of resistivity 8-12 Ω-cm were used as the substrates in all the experiments. Some of the wafers had 120 Å or 60 Å of thermal oxide grown at 850°C. A few wafers had the very thin oxide formed by boiling the wafers in the peroxide cleaning solutions. The other wafers were dipped in hydrofluoric acid after the cleaning treatment, and may therefore be expected to have the cleanest surfaces. The wafers then had 3000 Å of polysilicon deposited on them at 750°C at atmospheric pressure from silane with hydrogen as the diluent.

After the polysilicon deposition the samples were implanted with arsenic at 50 KeV to a dose of 10^{16} atoms/cm^2. 5000 Å of silicon dioxide were then deposited on the samples at 450°C. This oxide layer is to protect the surfaces of the samples from the furnace ambient and to prevent the evaporation of arsenic in the high temperature diffusions that were performed next. The annealings were carried out in a nitrogen ambient and varied from 1 hour at 1000°C to 7 days at 800°C.

Ellipsometric measurements on the sample which had oxide grown by boiling in hydrogen peroxide solutions, gave an oxide thickness of 20 Å. But Auger analysis indicated that the amount of oxygen in the layer is what would be present in 5 Å of stoichiometric silicon dioxide. Similarly the sample which received the hydrofluoric acid dip and is therefore nominally free of oxide was found to have the amount of oxygen that would be present in 2 Å of stoichiometric oxide.

Figs. 14.5(a) and 14.5(b) show the arsenic concentration profiles both in the polysilicon and in the single-crystal substrate for samples annealed at 1000°C for 4 hours. Fig. 14.5(a) is for the sample wich had "20 Å" of oxide between the polysilicon and the substrate while Fig. 14.5(b) is for the sample with "0 Å" of oxide. It is seen that the profiles are much the same for both the samples. There is a pile-up at the polysilicon/single-crystal silicon interface

which is a little less for the sample with "0 A" of oxide. The reason for this is that with the cleaner interface the substrate causes some of the grains in the polysilicon near the interface to be reoriented. This was also confirmed by channeling measurements. In the sample with "0 A" of oxide, the arsenic has diffused slightly deeper in the substrate than in the sample with "20 A" of oxide. Figures 14.6(a) and 14.6(b) are arsenic concentration profiles for diffusion from a polysilicon doping source diffused at 850°C for 3 days and 800°C for 7 days. The difference between the surface concentration in the substrate and the average concentration in the polysilicon is clearly seen. The profiles in the substrate are seen to be reasonably well fit by the Chebyshev polynomial expression given by Nakajima et. al. [14.9]. This expression has been derived on the assumption that the diffusivity of arsenic in single-crystal silicon is directly proportional to the concentration at high concentrations. Since the polynomial expression fits the concentration profiles in the substrate, the diffusion mechanisms in the single-crystal silicon must be unaffected by the presence of a deposited polysilicon layer.

It was found that 60 A or more of oxide between the substrate and the deposited layer completely blocked any diffusion of arsenic from the polysilicon to the substrate for an anneal at 1000°C for 4 hours.

The observed pile-up at the interfaces has been explained in terms of segregation of arsenic to the grain boundaries [14.10]. Consider two sections, one through the middle of the polysilicon layer and one at the bottom interface. The bottom section is a plane which lies on an interface between the polysilicon grains and single-crystal silicon. The section through the middle of the deposited layer could pass through an occasional grain boundary, but is mainly through the bulk of the grains. The backscattering measurement averages over the 1 mm square area of the incident beam and therefore if arsenic segregates to the interfaces of the grains, a pile-up would be expected.

352

Less pile-up of arsenic is evident in Fig. 14.5(b) because in this sample some of the grains at the interface have changed their orientation to that of the substrate because of the very clean interface. Hence the interface between the <100> oriented silicon and the randomly oriented crystallites is no longer a perfect plane. The pile up at the interface is therefore smeared out.

Mandurah et. al. [11] have used electrical measurements to determine the amount of dopant not segregated to the grain boudaries by assuming that dopant which is segregated to the grain boundaries is electrically inactive. The relationship that they have found between N_{GB}, the concentration of dopant per cm^3 at the grain boundaries and N_G, the dopant concentration in the grains is

$$\frac{N_{GB}}{N_G} = A \frac{Q_s}{N_{Si}} \exp\left(\frac{Q_o}{kT}\right) \tag{14.3}$$

When arsenic is the dopant, at a concentration level around 2×10^{20} atoms/cm^3, the value of the heat of segregation, Q_o, was found to be 0.41 eV. The value of the pre-exponential factor AQ_s/N_{Si} was 0.022 for a grain size of 1600 A. The heat of segregation is not dependent on grain size while the pre-exponential factor should be inversely proportional to the grain size.

In our experiments we have two effects which are a manifestation of arsenic segregation to the grain boundaries--the pile-up of arsenic at the polysilcon/single-crystal silicon interface and the difference between the average arsenic concentration in the polysilicon and the arsenic concentration at the surface in the single-crystal silicon. We have attempted to correlate our measurements to calculations form the values of Mandurah et. al.

Measurements of the amount of arsenic piled up at the polysilicon/single-crystal silicon interface which is analogous to the grain boundary have values of 4.1×10^{14}, 5×10^{14} and 5.8×10^{14} atoms/cm^2, respectively, for samples annealed at 1000°C for 4 hours, 850°C for three days, and 800°C for seven days. Assuming

that the grains are cubic, Q_{GB}, the number of arsenic atoms per square centimeter of grain boundary is given by

$$Q_{GB} = \frac{1}{3} N_{GB} L \qquad (14.4)$$

where L is the grain size. Using Eqs. (14.3) and (14.4), and knowing the grain size of 1600 A measured by Mandurah et. al. and the average concentration in the polysilicon, Q_{GB} was calculated. Measured and calculated values are compared in Table 14.1. It is seen that there is fairly good agreement between the values obtained from the two entirely different measurement techniques.

Measured and calculated values of the ratio of the surface concentration of arsenic in the single-crystal silicon to the average concentration in the polysilicon are given in Table 14.2. The calculated values are N_G/N where N_G is the concentration in the grains and N is the total average concentration in the polysilicon given by $N_{GB} + N_G$. These values were calculated using Eq. 4 but scaling the pre-exponential factor to account for the different grain sizes. The grain sizes are 3800 A, 1800 A, and 1300 A for the films annealed at 1000°C for 4 hours, 850°C for 3 days and 800°C for 7 days respectively. Again there is good agreement between measured and calculated values.

14.3 SILICIDE CONTACTS TO SINGLE-CRYSTAL SILICON

During the last few years, the use of refractory metal silicides has been heavily investigated to form interconnections in VLSI circuits. It has been observed that silicides of tungsten (W), molybdenum (Mo), and tantalum (Ta) and titanium (Ti) have reasonably good compatibility with IC fabrication technology [14.3,14.12-14.16]. They have fairly high conductivity, they can withstand all the chemicals normally encountered during the fabarication process, thermal oxidation of their silicides can be accomplished in oxygen and steam to produce a passivating layer of SiO_2, their contacts to shallow p-n junctions are

354

reliable, and fine lines can be etched by plasma etching these materials. This indicates that the silicides of W, Mo, Ta and Ti will be used as interconnection layers in a manner very similar to polysilicon.

As an interconnection layer, the silicides are required to make contacts to heavily doped single crystal silicon and polysilicon. The silicide layers are incorporated early in the process sequence and are subjected to several high temperature steps afterwards. This causes interdiffusion of dopants, which could alter the device behavior.

Mochizuki et. al. [14.3] annealed $MoSi_2$ contacts to heavily phoshporus-doped N^+ diffusions and observed that up to 700°C the contact resistance remained low; however, above 700°C it increased rapidly. By interposing a layer of N^+ polysilicon between the $MoSi_2$ and single crystal silicon the increase in the contact resistance did not take place until 1000°C. They were not able to determine the cause of this increase. Similar behavior was observed by Inoue et. al. [14.4]. They recognized that the problem was caused by dopant diffusion from the silicon to the silicide, causing depletion of donor impurities at the surface and hence resulting in increased contact resistance. To overcome this problem they doped the silicide heavily with phosphorus prior to the anneal and found that in this case no increase in contact resistance took place even after annealing up to 1100°C. Their results are shown in Fig. 14.7; obviously in this case the loss of dopant did not take place and therefore the contact resistance remained stable.

In order to investigate these changes in contact resistance one must first understand the basics of metal to silicon (M-S) contacts [14.17,14.18]. Because of the work function differences a potential barrier is formed when a metal is put in contact with a semiconductor as shown in Fig. 14.8. If an M-S contact has no potential barrier at the interface, a true ohmic contact results which has a linear current-voltage characteristic and its resistance is just that of the bulk

semiconductor. However, M-S contacts can appear ohmic for other reasons as well. If the potential barrier is so low that a negligible amount of voltage drops across the interface, again an ohmic behavior is obser- ᴪ (e.g., Au and Pt on p-type Si). However, practical ohmic contacts on n-type ᵤi are made by metals such as Al, Au, and Pt, which are known to form fairly high barriers (0.7 - 0.9 eV) on lightly doped Si. Since it is also known that the barrier height is essentially independent of the semiconductor doping concentration (except for image force lowering which should be less than 0.2 eV for $N_D \le 10^{20}$ cm^{-3}), it is clear that relatively high barriers (> 0.5 eV) exist in such contacts on heavily doped silicon. Therefore, these contacts are ohmic for a different reason.

When the doping concentration is high the depletion region width becomes small as shown in Fig. 14.8(a). Therefore, electrons can easily tunnel through the barrier, in addition to the thermionic emission process. This added component of current reduces the voltage drop across the contact such that most of the applied voltage is dropped across the semiconductor bulk. Thus the terminal current-voltage characteristic of such a contact is ohmic.

Fig. 14.8(a) depicts the situation where the semiconductor is lightly doped. In this case, the depletion width is so wide that the only way the electrons can get into the metal is by thermionic emission over the potential barrier ϕ_B.

Fig. 14.9 shows plots of specific contact resistance as a function of doping density in n and p type Si for different barrier heights [14.17]. It is evident that contact resistance increases sharply as the doping density decreases. It can also be seen that higher barrier materials give higher contact resistance and are more sensitive to the doping density in the silicon. It is obvious from Fig. 14.9 that to cause a factor of ten change in contact resistance such as shown in Fig. 14.7 only a factor of 2 to 3 change in doping density is required.

This change is needed only near the silicide/silicon interface just sufficient to cause a change in barrier thickness [14.17,14.18] and could be caused by interdiffusion and segregation of dopant at high temperatures.

Tsai et. al. [14.5] have observed considerable diffusion effects in a WSi_2^+/N polysilicon structure. They deposited WSi_2 on phosphorus doped polysilicon and annealed it with and without an oxide cap and found that phosphorus diffused very rapidly from the polysilicon into WSi_2 and into the ambient if the WSi_2 was not capped with SiO_2. A 1000°C anneal in their case lowered the phosphorus concentration in the polysilicon by about an order of magnitude. Pan et. al. [6] have recently studied this phenomenon in more detail. In their work the diffusion of dopants (B, P, and As), from the doped polysilicon through the silicide of a polycide during 800° or 1000°C annealing, was observed. The redistribution of boron in a polycide structure was found to be quite different from that of phosphorus and arsenic. The dopant concentration in the WSi_2 layer depends on the type of dopant, annealing conditions, film, and the stoichiometrics of the silicide. The diffusivity of the dopants was found to be more than three orders of magnitude higher than in single crystal silicon.

This rapid diffusion and redistribution can explain the changes in contact resistance [14.3,14.4]. If the silicides have to be used as a contact material the modeling of dopant diffusion in silicides and segregation and redistribution at the silicide/silicon interface is of extreme importance.

In our work, we seek to expand the modeling ability of SUPREM to include silicide to single crystal and polycide to single crystal contacts, in addition to doped polysilicon to single crystal contacts. Electrical properties and impurity profiles are being investigated in these structures. Silicides of tungsten and tantalum (WSi_2 and $TaSi_2$) are being used to form contacts with

357

doped silicon. In our future work the contact resistance will be measured as a function of high temperature processing. This will allow us to obtain important information from the point of view of process design and will point to cases of more interest which need to be examined in more detail.

In order to understand the redistribution of dopants across the polysilicon, polycide/silicon and silicide/silicon interfaces and their implication for electrical behavior of the contacts, we are investigating the ability of various silicides to dissolve dopants from nearby silicon in terms of both the thermodynamics of the process as well as the kinetics.

The first question to be answered is what is the concentration of dopant that can be accommodated by a silicide in contact with doped silicon at thermal equilibrium. In the dilute limit, we would expect the silicide concentraion C_M to be related to the silicon concentration C_S by a simple Boltzmann relationship

$$C_M = C_S \, e^{Q/kT} \tag{14.5}$$

where Q could be positive or negative and also contain temperature dependent terms. One goal of this part of the program is to measure Q for important silicides and dopants.

In addition to strictly the thermodynamics of the process, the kinetics will also be important to device processing. As a second part of this study, we are investigating the diffusion of dopants through silicides and possible segregation that might occur at the silicide/silicon interface.

Techniques being applied include Auger sputter profiling ASP, RBS and secondary ion mass spectrometry SIMS. These techniques are being used to obtain elemental concentration profiles for the mesurements necessary to answer the questions raised above. Finally, any OED effects observed when polycide or silicide contact structures are thermally oxidized are of interest. The combination of

358

these various contacting schemes will greatly enhance the modeling capability of SUPREM; for each of the materials, the objective is to develop quantitative models suitable for SUPREM implementation.

At the present time this work is just getting underway. Initial studies of WSi_2/Silicon contacts are in progress, specific results will be reported in the future.

REFERENCES

[14.1] See, for example, H. Fu, J. Manoliu and J. L. Moll, 1980 IEDM Tech. Digest, 140.

[14.2] T. H. Ning and R. D. Issac, IEEE Trans. Electron Dev., ED-27, 2051 (1980).

[14.3] T. Mochizuki, T. Tsujimaru, M. Kashiwagi and Y. Nishi, IEEE Trans. Elect. Dev., ED-27, 1631 (1980).

[14.4] S. Inoue, N. Toyokura, T. Nakamura and H. Ishikawa, Semiconductor Silicon 1981, The Electrochem. Soc., 81-5, 596 (1981).

[14.5] M. Y. Tsai, F. M. d'Heurle, C. S. Petersson and R. W. Johnson, "Properties of WSi_2 Film on Poly-Si," J. Appl. Phys., 52, 5350 (1981).

[14.6] P. Pan, N. Hsieh, H. J. Geipel, Jr., and G. J. Slusser, J. Appl. Phys., 53(4), 3059 (1982).

[14.7] J. D. Plummer et al., "Computer-Aided Design of Integrated Circuit Fabrication Processes for VLSI Devices," SEL Report TR DXG 501-81, July 1981.

[14.8] T. I. Kamins and T. R. Cass, "Structure of Chemically Deposited Polycrystalline-Silicon Films," Thin Solid Films, 16, 147, 1973.

[14.9] Y. Nakajima, S. Ohkawa snd Y. Fukukawa, "Simplified Expression for the Distribution of Diffused Impurity," Japan. J. Appl. Phys., 10, 162, 1971.

[14.10] B. Swaminathan, E. Demoulin, T. W. Sigmon, R. W. Dutton and R. Reif, "Segregation of Arsenic to the Grain Boundaries in Polycrystalline Silicon," J. Electrochem. Soc., 127, p. 2227, 1980.

[14.11] M. M. Mandurah, K. C. Saraswat, C. R. Helms and T. I. Kamins, "Dopant Segregation in Polycrystalline Silicon," J. Appl. Phys., 51, 5755, 1980.

[14.12] T. Mochizuki, K. Shibata, T. Inoue, and K. Ohuchi, "A New MOS Process Using MoSi$_2$ as a Gate Material," Japan. J. Appl. Phys., Suppl., 17, 37-42, 1977.

[14.13] B. L. Crowder and S. Zirinsky, "1-μm MOSFET VLSI Technology. Part VII: Metal Silicide Interconnection Technology - Future Perspective," IEEE Trans. Electron Devices, Vol. ED-26, 369-371, 1979.

[14.14] F. Mohammadi and K. C. Saraswat, "Properties of Sputtered Tungsten Silicide for MOS Integrated Circuit Applications," J. Electrochem. Soc., 127, 450-454, 1980.

[14.15] S. P. Murarka, D. B. Fraser, A. K. Sinha, and H. J. Levinstein, "Refractory Silicides of Titanium for Low-Resistivity Gates and Interconnects," IEEE Trans. Elect. Dev., ED-27, 1425-1430, 1980.

[14.16] H. J. Geipel, J.r., N. Hsieh, M. H. Ishaq, C. W. Koburger, and F. R. White, "Composite Silicide Gate Electrodes-Interconnections for VLSI Device Technologies," IEEE Trans. Elect. Dev., ED-27, 1417-1424, 1980.

[14.17] C. Y. Chang, Y. K. Fang and S. M. Sze, Solid-State Electronics, 14, 541 (1971).

[14.18] A. Y. C. Yu, Solid State Electronics, 13, 239 (1970).

Table 14.1

Calculated and Measured Arsenic Concentrations in the Grain Boundary

atoms/cm^2 of grain boundary		
Temperature (°C)	Measured	Calculated
1000	4.1×10^{14}	4.4×10^{14}
850	5.0×10^{14}	7.1×10^{14}
800	5.8×10^{14}	8.3×10^{14}

Table 14.2

Ratio of Surface Concentration to Poly Concentration

$C_{surface}/C_{poly}$		
Temperature (°C)	Measured	Calculated
1000	0.75	0.71
850	0.41	0.42
800	0.24	0.30
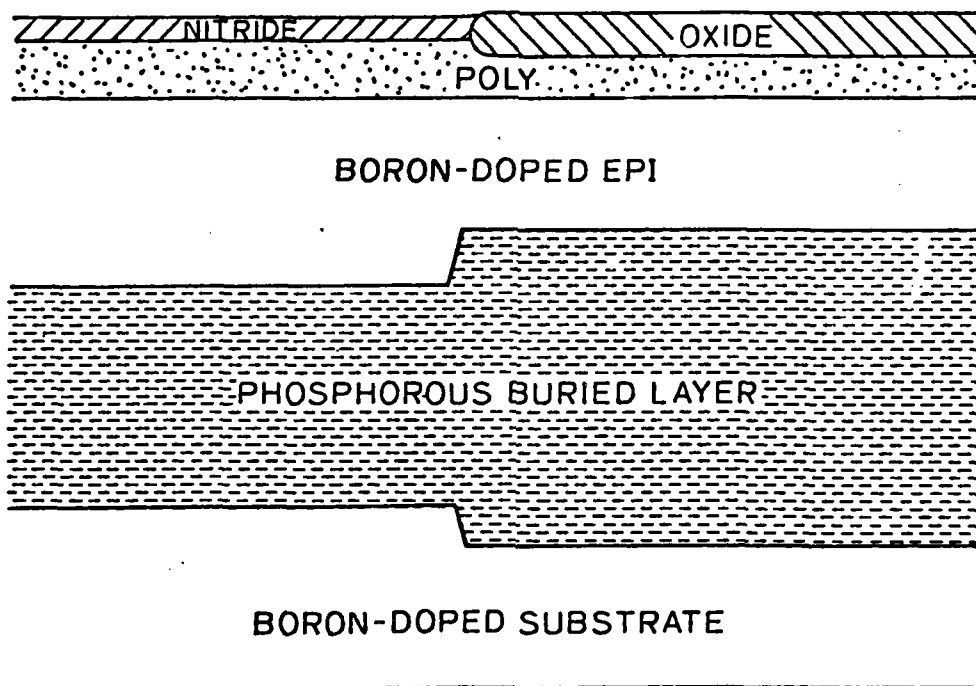
Fig. 14.1: Structure used to demonstrate the absorption of interstitials by the grain boundaries in polysilicon.
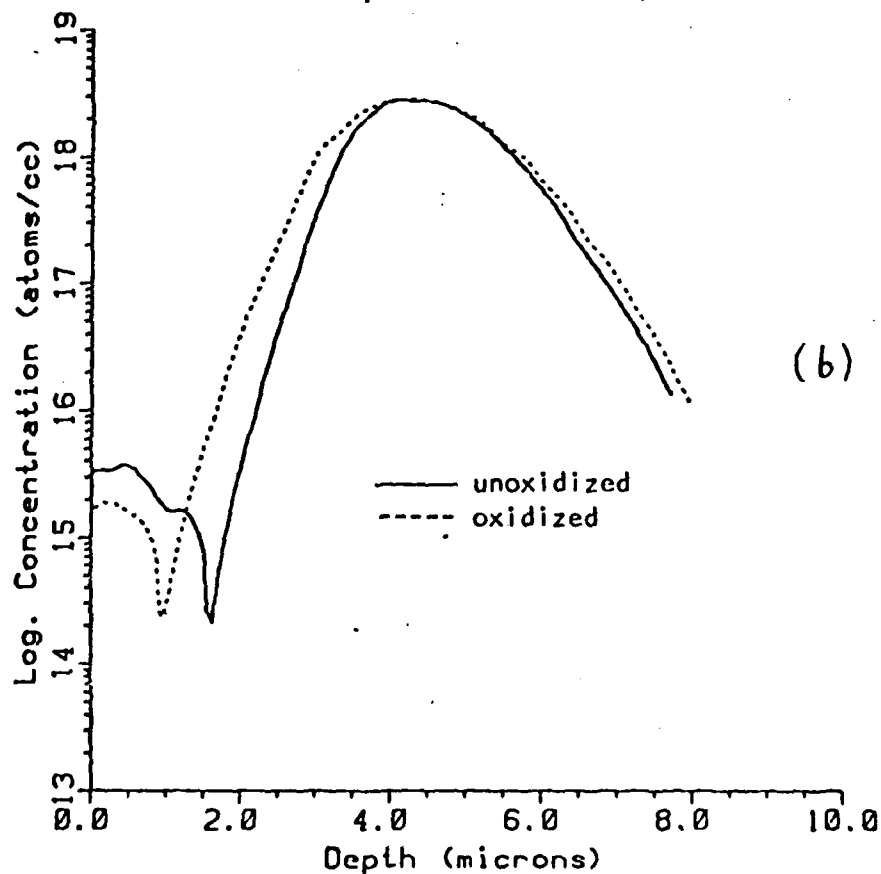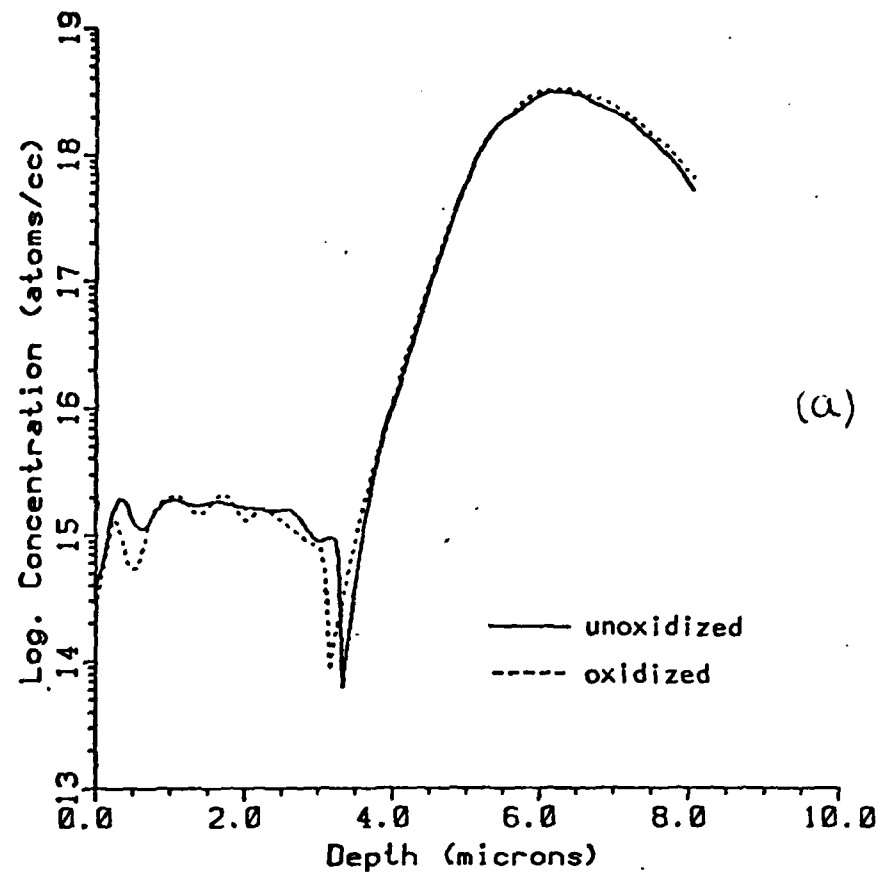
Fig. 14.2: Buried layer concentration profiles in the structure of Fig. 14.1 after diffusion in dry O_2 for 17.5 hours. The polysilicon thicknesses are 3000 Å in (a) and 6000 Å in (b).

SINGLE CRYSTAL SILICON SUBSTRATE
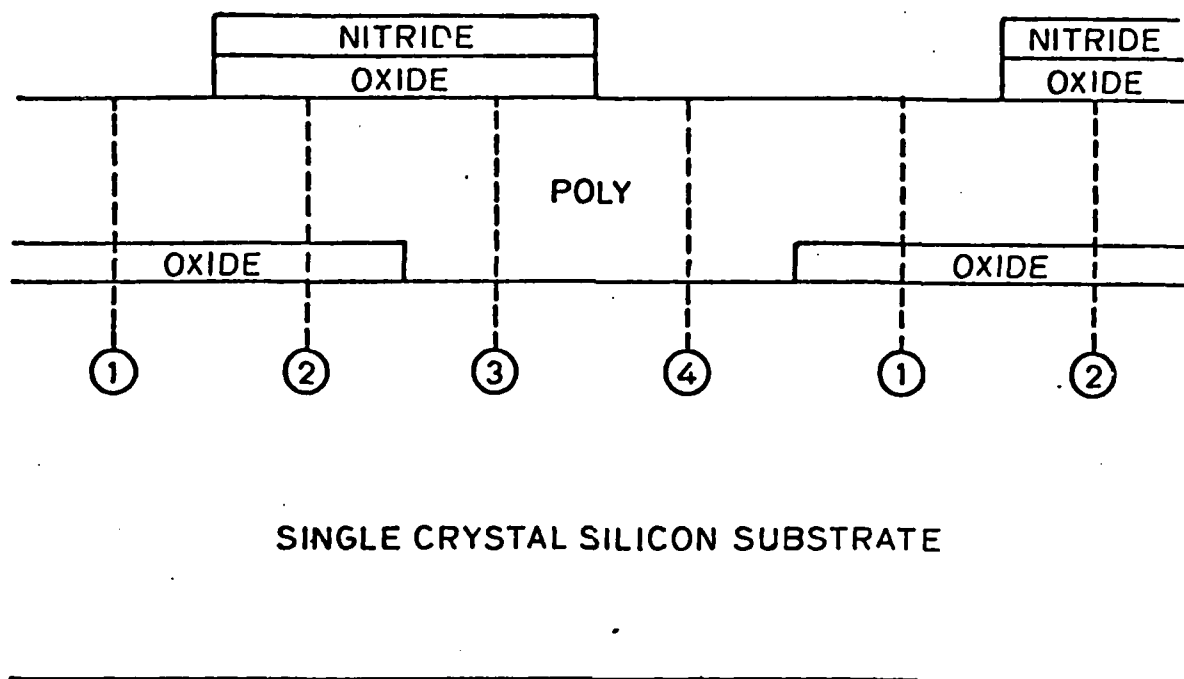
Fig. 14.3: Structure used in experiments on the effect of polysilicon grain size on OED.
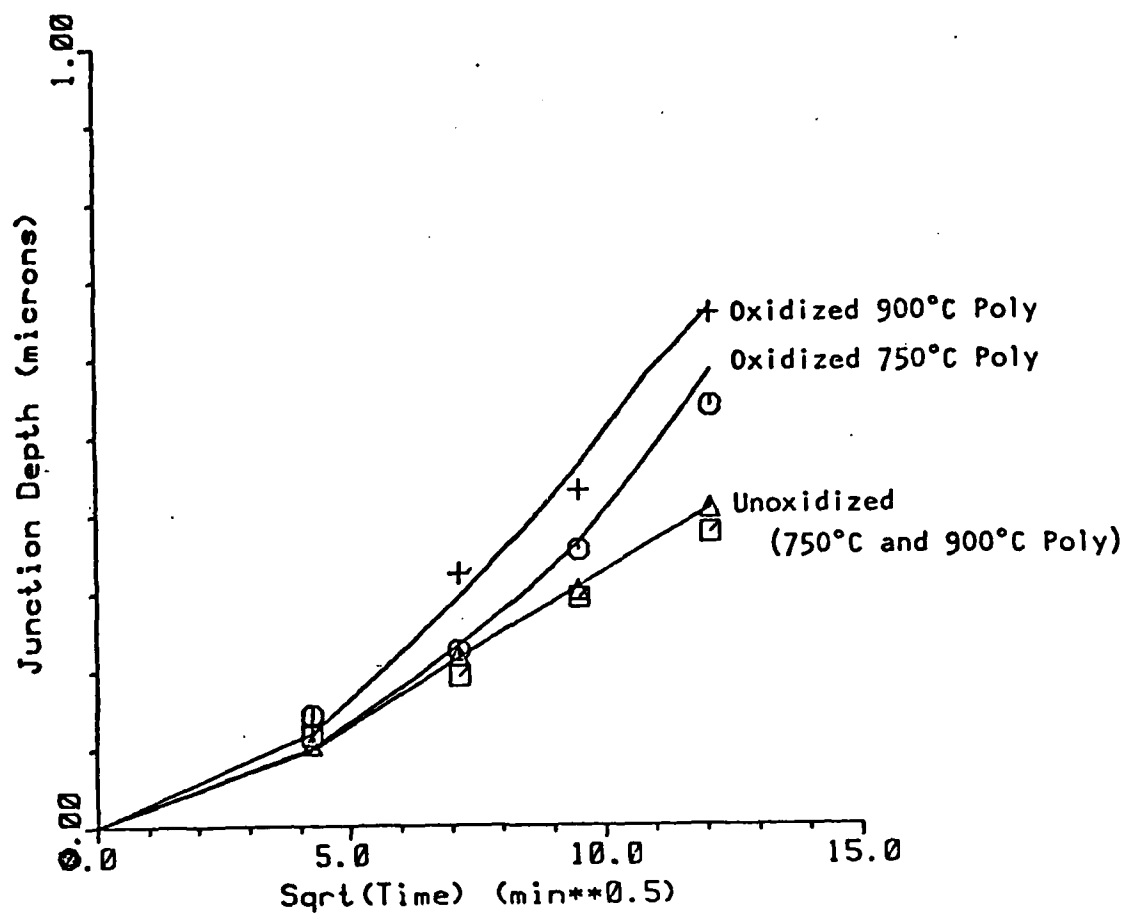
Fig. 14.4: Junction depth vs. $\sqrt{t}$ for samples with 3000 Å of polysilicon deposited at 750°C and 900°C. The junction depth is greater for the sample depostied at 900°C in the oxidized region but the same in the unoxidized region.
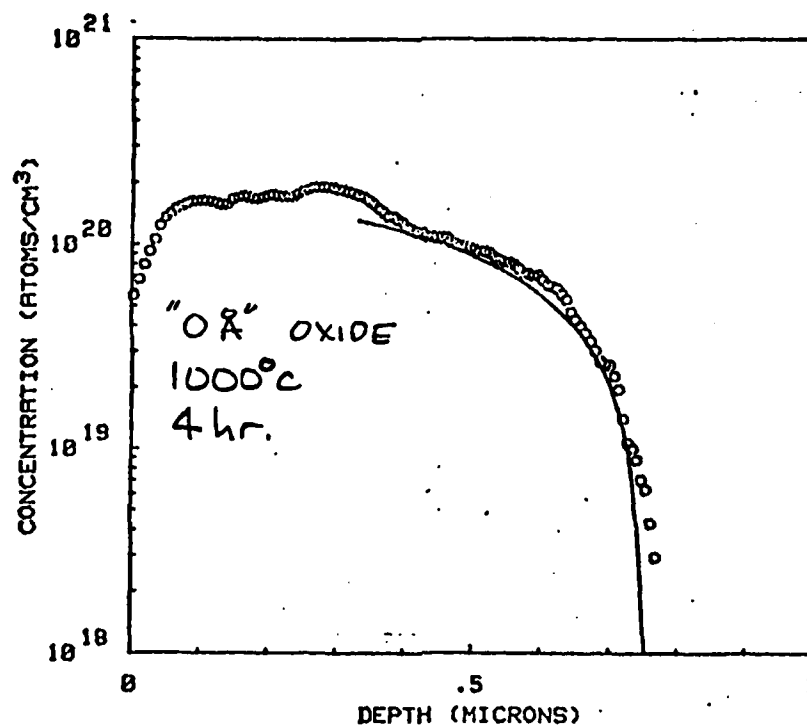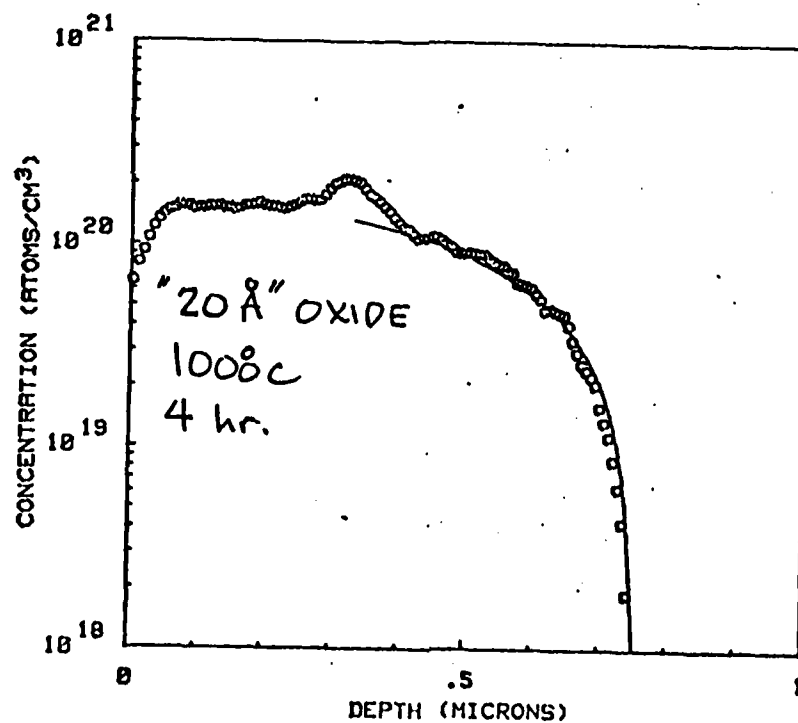
Fig. 14.5: Concentration profiles after diffusion of arsenic from a 3000 Å
polysilicon doping source at 1000°C for 4 hours.
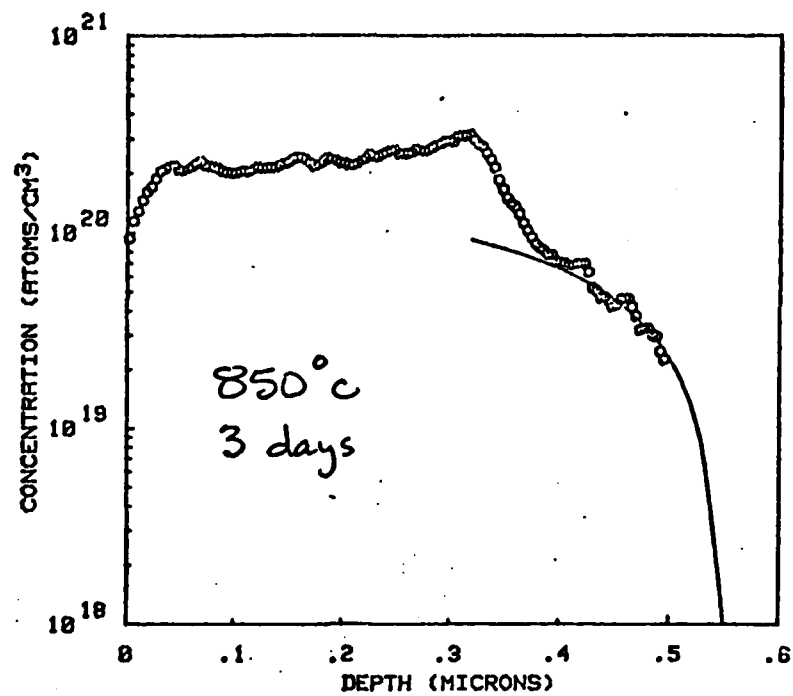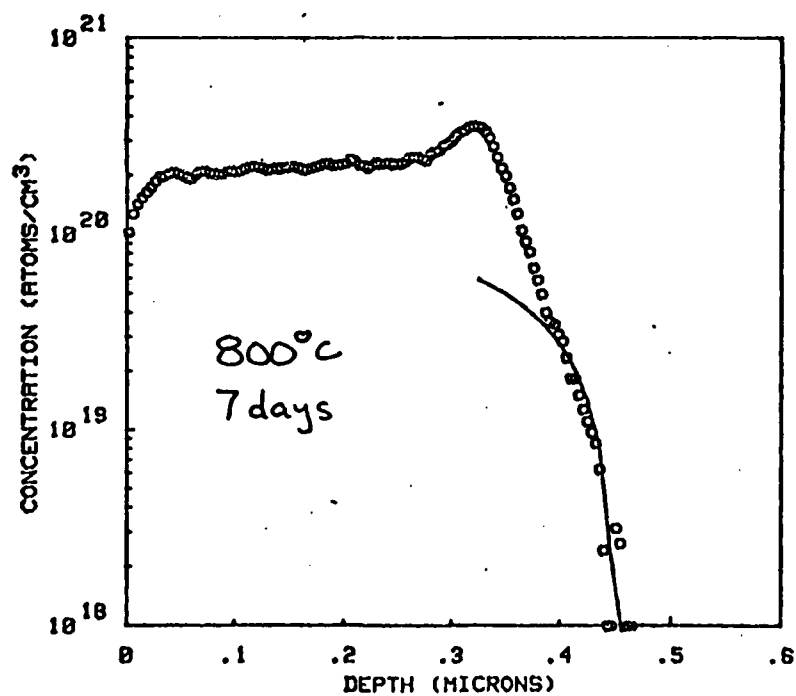
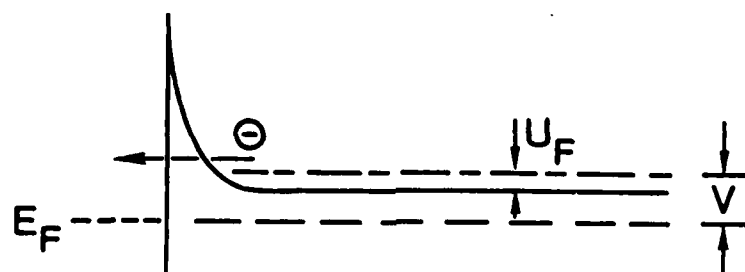Fig. 14.6: Arsenic profiles after diffusion from a polysilicon doping source for (a) 850°C, 3 days and (b) 800°C, 7 days.

367
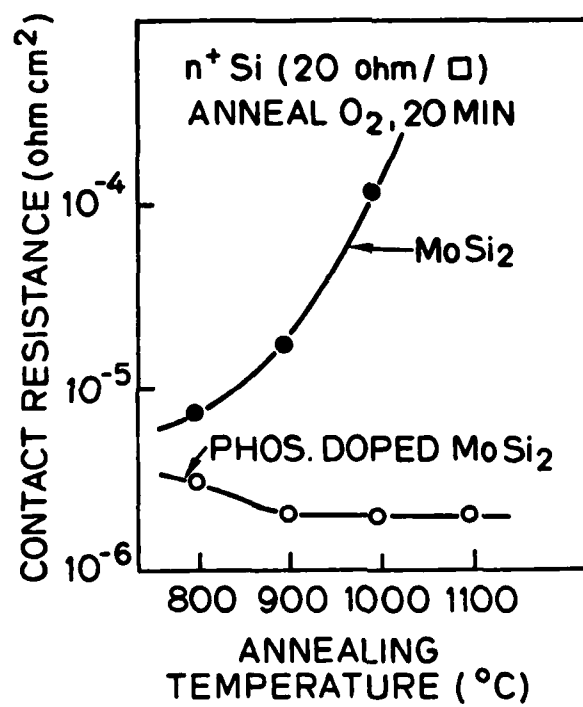
Fig. 14.7: Contact resistance between Silicide and N$^+$ Si after annealing for 20 min. [14.4].
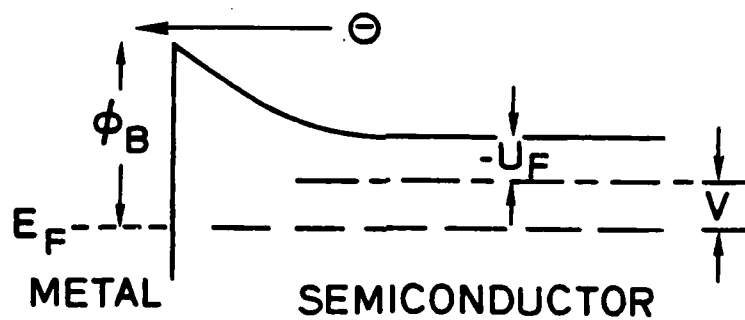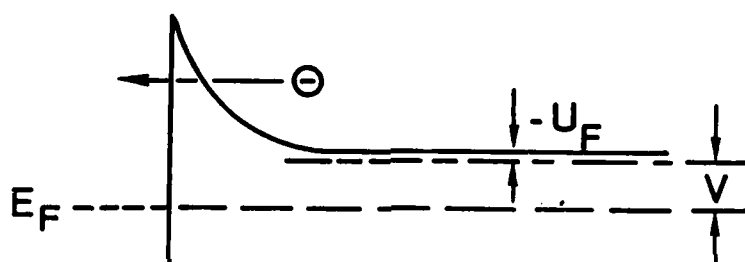
Fig. 14.8: Band diagrams of metal n-type semiconductor contacts under forward bias voltage: (a) semiconductor lightly doped, (b) semiconductor heavily doped, (c) semiconductor very heavily doped (degenerate).
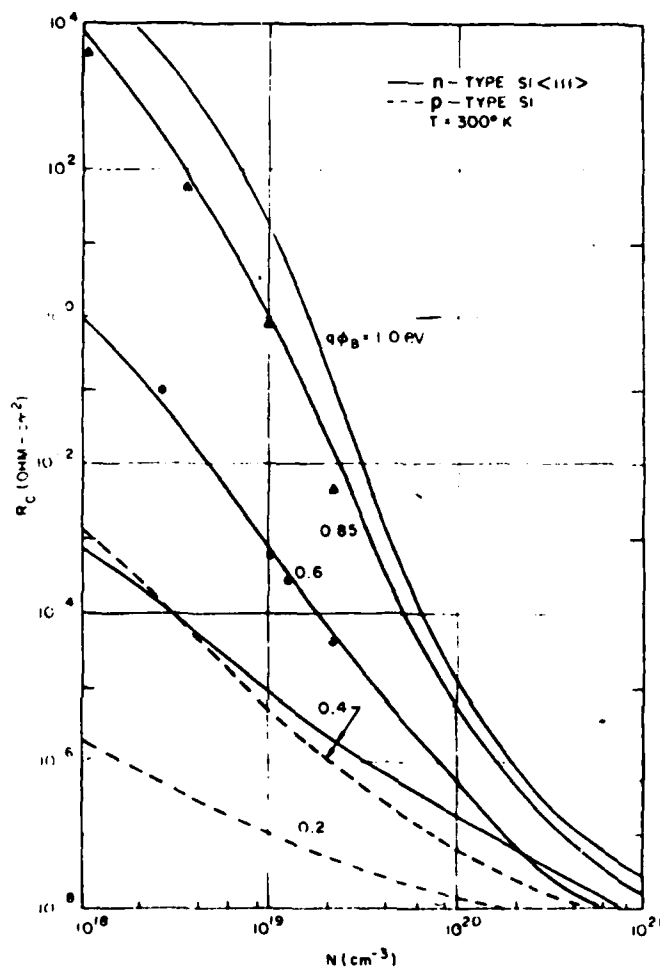
Fig. 14.9: Theoretical specific contact rsistance at 300°C for n-type. <111> oriented (solid lines) and p-type (dotted lines) silicon. The solid circles are experimental results of Al-Si and Mo-Si barriers ($q\psi_B$ = 0.6 eV), and triangles for PtSi-Si barriers ($q\psi_B$ = 0.85 eV) [14.17].

15. DIFFUSION OF ARSENIC IN POLYCRYSTALLINE SILICON

B. Swaminathan, M. Kump, K. Saraswat, R. Dutton

15.1 INTRODUCTION

The results of studies on the diffusion of dopants from polysilicon doping sources to the single-crystal silicon substrate in various ambients have been detailed in another section of this report. In this application, the value of the diffusivity of the dopant in the polysilicon itself is usually of little consequence, since the diffusivity is so large that the dopant is distributed uniformly in the polysilicon, and the overall profile is determined by the diffusivity in the single-crystal silicon. However the characterization of dopant diffusion in polysilicon is of interest for some device applications. The performance of high value polysilicon resistors used in static memory is limited by the diffusion of dopant, primarily along the grain boundaries, from the highly doped contact area to the lightly doped resistor area. Diffusion along the grain boundaries will also limit the performance of devices fabricated on large-grained polysilicon thin films obtained by laser recrystallization.

The diffusion of arsenic in polycrystalline silicon films has been studied over the temperature range of 750 to 950°C and for grain sizes from 0.21 μm to 0.51 μm [15.1]. Rutherford backscattering spectrometry has been used to measure the concentration profiles of arsenic, initially introduced into the polysilicon by ion implantation, after various annealing steps. The concentration profiles have been found to be determined by a combination of several factors--a low diffusivity in the bulk of the grains, a much higher diffusivity in the grain boundaries and the segregation of arsenic to the grain boundaries. The diffusivity of arsenic in the grain boundaries is independent of concentration, with an activation energy of 3.9 eV, very close to that of the low-concentration arsenic diffu-

371

sivity in single-crystal silicon. However, the value of the diffusivity is 1.3 x 10^5 exp (-3.9/kT) cm^2/sec, four orders of magnitude higher than the single-crystal value. The diffusivity in the interior of the grains is the same as that in single-crystal silicon.

15.2 EXPERIMENTAL PROCEDURE

<100>-oriented silicon wafers with 1000 A of thermally grown silicon dioxide were used as substrates for the deposition of 1 μm thick layers of polysilicon in a low-pressure chemical-vapor-deposition reactor at the pressure of 0.2 Torr and a temperature of 625°C. To study the effect of changing grain size and structure on the diffusivity of arsenic, some of the polysilicon films were annealed in a nitrogen ambient at 1100°C or 1200°C for one hour. Transmission electron micro-scopy was used to measure the grain sizes of the three kinds of films. The as-deposited films had an average grain size of 0.21 μm, the films annealed at 1100°C had a grain size of 0.36 μm, and the films annealed at 1200°C had a grain size of 0.51 μm. X-ray diffraction was used to measure the texture of the var-ious films. The as-deposited films had a great deal of <110> texture and very little of any other texture. The <110> texture was less for the films annealed at 1100°C and still less for the films annealed at 1200°C. The other textures, particularly <311>, increased with higher temperature annealing. These results are consistent with the work of Kamins, et. al. [15.2].

The three kinds of films were implanted with 40 KeV arsenic ions to a dose of 10^{16} atoms/cm^2, and a 4000 A layer of silicon dioxide was chemically vapor deposited on the films at 450°C. The samples were then subjected to heat treat-ments in a nitrogen ambient, ranging from 3 minutes at 950°C to 24 hours at 750°C. For the shortest diffusion times, heating transients were minimized by pushing the wafers into the furnace without a boat. X-ray diffraction confirmed that these heat treatments were for such short times or were at such low tempera-

372

tures that little grain growth or change in structure of the polycrystalline silicon could occur. The arsenic distributions in the samples before and after the heat treatments were measured using 2.2 meV, $^4He^+$ ion backscattering.

15.3 EXPERIMENTAL RESULTS AND INTERPRETATION

Fig. 15.1 shows the concentration of samples annealed at 800°C for six hours. It is seen that there is a sharp peak in each profile near the surface of the polysilicon layer. The concentration then drops off more slowly with depth. This latter portion of the profile is well fitted with a complementary error function as shown.

These concentration profiles can be explained by the low diffusivity of the arsenic in the interior of the grains and the much higher diffusivity in the grain boundaries. After the 40 KeV ion implantation, the arsenic is mostly in the interior of the grains in a region less than 1000 A from the surface. On heat treatment, some of the arsenic moves from the bulk of the grains to the grain boundaries, where it tends to segregate [15.3,15.4]. The amount of arsenic reaching the grain boundaries is limited by the slow diffusivity in the bulk of the grains and arsenic remaining in the interior of the grains is responsible for the peak near the surface in the concentration profiles of Fig. 15.1. The arsenic which reaches the grain boundaries diffuses very rapidly there. This source-limited diffusion tends to produce a complementary error function concentration profile, similar to the case of diffusion from a low-diffusivity layer into a substrate with much higher diffusivity [15.5].

Since the portions of the concentration profiles resulting from grain-boundary diffusion are well fitted by complementary error functions, the grain-boundary diffusivity must be independent of concentration. This is in contrast to the case for diffusion in single-crystal silicon where there is a strong con-centration dependence of the diffusivity because of which the profile deviates

from analytic forms such as the Gaussian or the complementary error function. Values for the grain-boundary diffusivity were extracted from the concentration profiles using the parameters found by fitting complementary error functions to the data. For the samples shown in Fig. 15.1, the grain-boundary diffusivity is $3.7 \times 10^{-14} cm^2$/sec at 800°C. This value is about the same for the initially unannealed polysilicon sample of 0.21 μm grain size, and for the samples with grain sizes of 0.36 μm and 0.51 μm.

Fig. 15.2 shows an Arrhenius plot of the measured grain-boundary diffusivity. The data are seen to be described very well by 1.3×10^5 exp $(-3.9/kT)$ cm^2/sec. The diffusivity of arsenic in the polysilicon grain boundaries is four orders of magnitude higher than the diffusivity at low concentrations in single-crystal silicon, but the activation energy is quite close to the single-crystal value of 4.1 eV [15.6]. These results are surprising since the activation energy of diffusion in the grain boundaries of metals is only half that of diffusion in the bulk [15.7].

It is seen in Fig. 15.1 that the larger-grained polysilicon samples have larger peaks near the surface and less arsenic diffused deep into the samples. In the larger-grained material, the arsenic has to diffuse a greater distance in the grain before reaching a grain boundary, hence more of the arsenic remains in the interior of the grains.

These measurements of the concentration profile do not readily yield information on the value of the diffusivity in the bulk of the grains. Still a rough estimate can be made from the concentration profile of Fig. 15.3, which is for a sample of polysilicon with 0.51 μm grains, implanted with arsenic and annealed at 900°C for 1 hour. The arsenic which has diffused out to the grain boundaries is distributed uniformly through the depth of the film, while the arsenic still in the interior of the grains near the surface forms a peak in the concentration

profile. This peak is broader than the initial ion-implanted concentration
profile because of bulk diffusion. The broadened surface peak may not be Gauss-
ian because of concentration-dependent diffusion. Considering the limited exper-
imental depth resolution of about 400 A, an approximate diffusivity may be
extracted by fitting the surface peak with a Gaussian distribution. The
resulting value for the bulk diffusivity is 8×10^{-15} cm^2/sec at 900°C, compared
to a grain-boundary diffusivity of 1.6×10^{-12} cm^2/sec at the same temperature.
Diffusivity in single-crystal silicon of arsenic at a concentration of 2×10^{20}
atoms/cm^3 is also about 8×10^{-15} cm^2/sec. Thus the diffusivity in the interior
of the grains seems to be the same as that in single crystal silicon.

We used the values provided by Mandurah et al. [15.3] for arsenic
grain-boundary segregation, to verify that most of the arsenic which has left the
surface implanted layer remains in the grain boundaries. For the heat treatments
used in our experiments, the amount of arsenic transported by lattice diffusion
away from the grain boundaries into the grains was small compared to the amount
in the grain boundaries, and the situation corresponds to Harrison's "type C"
[15.8]. The simple analysis which fits the measured profiles by complementary
error functions is therefore justified, and it is not necessary to use the much
more involved analysis of Whipple [15.9].

15.4 2-D NUMERICAL SIMULATION OF DIFFUSION IN A GRAIN

The two-dimensional process simulation program SUPRA [15.10] has been used
to simulate the diffusion of arsenic in a grain of polysilicon with a grain
boundary. The polysilicon layer is idealized as having equally spaced grain
boundaries perpendicular to the surface as shown in Fig. 15.4(a). There is
reflection symmetry about the center of a grain between grain boundaries and
about the middle of a grain boundary, for initial conditions consisting of a sur-
face implanted layer of dopant. Therefore only one half of a grain needs to be

considered. The structure used in the simulation is shown in Fig. 15.4(b). The grain boundary region is modeled by a 5 A wide sliver at the right edge of the structure, while the rest of the structure simulates the interior of a grain.

The grain boundary region is specified to have a dopant diffusivity independent of concentration and a value much higher than in the interior of the grain as indicated by the experimental results discussed previously. The segregation of arsenic to the grain boundary is modeled by specifying a segregation coefficient. Assuming cubic grains, the segregation coefficient, m, is given by

$$m\delta = \frac{1}{6} \frac{N_{GB}}{N_G} L \qquad (15.1)$$

where 2δ is the thickness of the grain boundary, N_{GB}/N_G is the ratio of the concentration of arsenic in the grain boundary to that in the interior of the grain and may be calculated from the values given by Mandurah et al. [15.3], and L is the grain size. At 800°C $m\delta$ is approximately 500 A.

As discussed above, the diffusivity in the interior of the grains seems to be about the same as in single-crystal silicon. In the simulations the dependence of the diffusivity on the concentration as given by Fair [15.11] were used. Both the increase in the diffusivity at high concentrations because of the increase in the vacancy concentration, as well as the decrease in the diffusivity caused by arsenic clustering are taken into account.

The results of simulation for a sample implanted with 10^{16} atoms/cm^2 of arsenic and annealed at 800°C for 6 hours are shown in Fig. 15.5. The contours of constant concentration after redistribution are shown in Fig. 15.5(a). It is seen that the arsenic at the top of the vicinity of the grain boundary is depleted. The arsenic diffuses rapidly down the grain boundary and back from the grain boundary to the interior of the grain. At any depth greater than 0.2 μm, less than 10% of the dopant is in the interior of the grain, most of the dopant is in

the grain boundary. In order to compare the simulation with experimental results, the two-dimensional concentration profiles had first to be reduced to the one-dimensional concentration profile obtained from the RBS measurement which gives the average concentration at any horizontal plane. That is

$$C_1(y) = \frac{1}{W} \int_0^W C_2(x,y)dx \qquad (15.2)$$

where $C_1(y)$ is the effective 1-D concentration profile, W is the width of the simulation region and $C_2(x,y)$ is the 2-D concentration profile. In addition, one other effect has to be taken into account. The diffusion in a grain with grain boundaries is actually a three-dimensional problem, but the idealized structure on which the simulation has been performed, Fig. 15.4(a), is a two-dimensional one. The grain boundaries should run not only perpendicular to the plane of the paper as in Fig. 15.4(a), but also parallel to the plane of the paper and with the same spacing as the other set of grain boundaries, if grains with a square cross section are assumed. This effect can be corrected for by weighing the regions near the grain boundary by a factor of 2 in taking the section average indicated in Eq. 15.2. The comparison of experiment with simulation is shown in Fig. 15.5(b). The grain size and the diffusivity in the grain boundary have been used as adjustable parameters to fit the experimental data. The value of the grain size used in the simulations is half the value of 3600 A measured by transmission electron microscopy. This discrepancy might be due to uncertainties in the procedure for measuring the grain size from a transmission electron micrograph, and to errors involved in assuming grains with a square cross section in correcting for the three dimensional effects. The value of the diffusivity in the grain boundary used in the simulation is 4.2×10^{-14} cm^2/sec as compared to 3.8×10^{-14} cm^2/sec, the value extracted by fitting the data with a complementary error function. This difference is because of the small amount of arsenic which has diffused back into the grain from the grain boundary.

377

For the conditions in which the experiments with arsenic were performed, the interpretation of the profiles is much simplified because the profile is neatly separated in two regions, a surface region where the dopant is mostly in the interior of the grains and a tail region where the dopant is mostly in the grain boundary. This circumstance occurs because of the large amount of arsenic segregation to the grain boundary. The work of Mandurah et al. [15.3] has shown that in the case of boron there is no segregation to the grain boundary, and hence most of the dopant is always in the interior of the grain. Fig. 15.6 shows the results of simulation for boron diffusion with a grain boundary diffusivity 10^4 times the diffusivity in the interior of the grain. It is seen that the profile follows no simple pattern.

The amount of lateral diffusion that might be expected from a contact region in a polysilicon resistor or transistor structure has been simulated. Fig. 15.7 is the top view of a polysilicon layer with a highly doped arsenic region on the le1c and with a grain boundary. After an inert ambient anneal for 5 hours at 1000°C, the arsenic has segregated to the boundary and diffused along it to produce a finger of approximately 5 μm length. This compares well with EBIC measurements of the same structure by Johnson et al. [15.14].

15.5 DIFFUSION MECHANISMS IN THE GRAIN BOUNDARY

The arsenic concentration profile in the grain boundary has been shown above to be of the form $N_0 erfc(y/2\sqrt{Dt})$ where D is the diffusivity in the grain boundary and t is the diffusion time. N_0, the surface concentration in the grain boundary is determined by the amount of arsenic implanted in the polysilicon, and by the concentration-dependent diffusivity in the interior of the grain. The grain boundary diffusivity is independent of concentration, and since the activation energy is 3.9 eV, very close to that of the diffusivity at low concentrations in single-crystal silicon, it may be suspected that the diffusion mechanisms are similar in the grain boundary and in intrinsic silicon.

In the case of metals the activation energy for diffusion in the grain boundaries is often only half that for diffusion in the bulk [15.7]. This might be because of the great difference between the neighborhood of a diffusion atom in the interior or a grain and that of one in a grain boundary. The interior of a grain in a metal usually has the atoms close-packed around the diffusing atom, but at the disordered region constituting a grain boundary, the packing is much looser. However for the case of silicon, because of the very open diamond lattice, the structure of the immediate neighborhood around an atom may be similar whether it is in the interior of a grain or in a grain boundary. This could account for the activation energies of diffusion in bulk silicon and in the grain boundary being so close.

Impurity and self-diffusion in single-crystal silicon are believed to occur predominantly via interactions with vacancies in various charge states. The concentrations of the charged vacancies are a function of the Fermi level. The activation energy for impurity diffusion with a vacancy in the rth charge state is [15.11]

$$Q_I^r = H_V^r + Q_{IV}^r - E_b^r$$

where (15.3)

$$Q_{IV}^r = \Delta H_m^r - \Delta Q^r + E_b^r$$

Here H_V^r is the enthalpy of formation of the vacancy, Q_I is the effective activation energy for diffusion of the impurity with the vacancy, Q_{IV}^r is the migration energy of the impurity-vacancy complex, E_b^r is the binding energy of the impurity-vacancy pair, ΔH_m^r is the migration enthalpy of the vacancy and ΔQ^r is the potential energy difference between a vacancy at a third coordination site from the impurity and one at infinity.

For arsenic in silicon Fair [15.11] has given

$$D_{As} = 0.066 \exp \left(- \frac{3.44}{kT} \right) + 12.0 \left(\frac{n}{n_i} \right) \exp \left(- \frac{4.05}{kT} \right) \text{ cm}^2/\text{sec} \qquad (15.4)$$

where 3.44 eV is the activation energy for diffusion with neutral vacancies and 4.05 eV is the activation energy for diffusion with negatively charged vacancies. We estimate from our data that the activation energy of grain boundary diffusivity of arsenic is 3.9 ±0.2 eV. Comparing this to Fair's data might lead one to conclude that the diffusion in the grain boundary is with negatively charged vacancies. However it would not be unexpected that the activation energies for diffusion in the bulk and in the grain boundary be somewhat different. Moreover the data of Fair are questionable. Fair in an earlier paper [15.12] gives an expression for the diffusivity of arsenic in intrinsic silicon which is close to the second term of Eq. 4 and which is also close to the data of Chiu and Ghosh [15.6]. The earlier work is referred to in a later paper [15.13] but the diffusivity given earlier is stated to be only the component from negatively charged vacancies, and a contribution from neutral vacancies is added. This is a substantial term and modifies the diffusivity in intrinsic silicon by a factor of 2 or more at all temperatures below 1150°C. No reason for this discrepancy with his own earlier data or with the data of others is given. Drawing definitive conclusions based on these results would therefore be unwarranted.

The best explanation for the independence of the diffusivity from the concentration level in the grain boundary would be that the Fermi level in the grain boundary is not a function of the arsenic concentration because the arsenic in the grain boundary is not electrically active. In the case of diffusion in single-crystal silicon, it is the dependence of vacancy concentration on the Fermi level which causes most of the concentration dependence of the diffusivity. The diffusion mechanisms might therefore be about the same in the grain boundary as in intrinsic silicon, but with a much higher concentration of vacancies.

15.6 COMPARISON WITH PREVIOUS WORK

There are differences between our results and published work on diffusion on polysilicon. Kamins et al. [15.15] reported a strong correlation between <110> texture and the diffusivity of boron and phosphorus in thick films of polysilicon. For the thinner films used in this work, no such relationship was found, the arsenic grain boundary diffusivity being independent of X-ray texture.

The grain-boundary diffusivities of arsenic obtained in this work are two orders of magnitude higher than the values reported by Tsukamoto et al. [15.16] and Ryssel et al. [15.17]. We believe that this difference is caused by an incorrect interpretation of the data in their work. As discussed above, the broadening of the surface peak in Fig. 15.3, gives an indication of the diffusivity in the interior of the grains. This value of diffusivity is the same as in single-crystal silicon at the same dopant concentration, and is two orders of magnitude lower than grain boundary diffusivity. It appears that Tsukamoto et al. and Ryssel et al. have extracted all their diffusivity values from measurements of the surface peak.

Johnson et al. [15.14] have reported results of electron-beam-induced-current and transmission-electron-microscope measurements on arsenic-diffused lateral diodes fabricated on laser-recrystallized polysilicon diodes. As mentioned above their results are consistent with our data.

15.7 SUMMARY

The diffusivity of arsenic in the grain boundaries in polycrystalline silicon has been studied experimentally and modeled with two-dimensional process simulation. It has been found that the diffusivity is much higher than reported by earlier workers. The activation energy for diffusion in the grain boundary is close to that for diffusion in the bulk. The diffusion in the interior of the grain is the same as in single-crystal silicon. Simulations have shown that the

concentration profiles in polycrystalline silicon after diffusion are very much influenced by the amount of segregation to the grain boundaries. Boron profiles are quite different from arsenic ones because there is no segregation of boron to the grain boundary.

REFERENCES

[15.1] B. Swaminathan, K. C. Saraswat, R. W. Dutton and T. I. Kamins, "Diffusion of Arsenic in Polycrystalline Silicon," Appl. Phys. Lett., 40, p. 795, 1982.

[15.2] T. I. Kamins, M. M. Mandurah and K. C. Saraswat, "Structure and Stability of Low Pressure Chemically Vapor Deposited Silicon Films," J. Electrochem. Soc., 125, p. 927, 1978.

[15.3] M. M. Mandurah, K. C. Saraswat, and T. I. Kamins, "Dopant Segregation in Polycrystalline Silicon," J. Appl. Phys., 51, p. 5755, 1980.

[15.4] B. Swaminathan, E. Demoulin, T. W. Sigmon, R. W. Dutton and R. Reif, "Segregation of Arsenic to the Grain Boundaries in Polycrystalline Silicon," J. Electrochem. Soc., 127, p. 2227, 1980.

[15.5] M. L. Barry and P. Olofsen, "Doped Oxides as Diffusion Sources: I. Boron into Silicon," J. Electrochem. Soc., 116, p. 2227, 1980.

[15.6] T. I. Chiu and H. N. Ghosh, "A Diffusion Model for Arsenic in Silicon," IBM J. Res. Dev., 15, p. 472, 1971.

[15.7] N. A. Gjostein in "Diffusion," Am. Soc. Metals, Metals Park, Ohio, p. 241, 1973.

[15.8] L. G. Harrison, "Influence of Dislocations on Diffusion Kinetics in Solids with Particular Reference to the Alkali Halides," Trans. Faraday Soc., 57, p. 1191, 1961.

[15.9] R. T. P. Whipple, "Concentration Contours in Grain Boundary Diffusion," Phil. Mag., 45, p. 1225, 1954.

[15.10] D. Chin, M. Kump, and R. W. Dutton, "SUPRA: Stanford University Process Analysis Program," Stanford, California, 1981.

[15.11] R. B. Fair in "Impurity Doping Processes in Silicon," edited by F. Y. Wang, North Holland Publishing, Amsterdam, 1981.

[15.12] R. B. Fair and J. C. C. Tsai, "The Diffusion of Ion-Implanted Arsenic in Silicon," J. Electrochem. Soc., 122, p. 1689, 1975.

[15.13] R. B. Fair and J. C. C. Tsai, "A Quantitative Model for the Diffusion of Phoshporus in Silicon and the Emitter Dop Effect," *J. Electrochem. Soc.*, 124, p. 1107, 1977.

[15.14] N. M. Johnson, D. K. Biegelsen and M. D. Moyer, "Grain Boundaries in P-N Junction Diodes Fabricated in Laser-Recrystallized Silicon Thin Films," *Appl. Phys. Lett.*, 38, p. 900, 1981.

[15.15] T. I. Kamins, J. Manoliu, and R. N. Tucker, "Diffusion of Impurities in Polyrystalline Silicon," *J. Appl. Phys.*, 43, p. 83, 1972.

[15.16] K. Tsukamoto, Y. Akasaka and K. Horie, "Arsenic Implantation into Polycrystalline Silicon and Diffusion to Silicon Substrate," *J. Appl. Phys.*, 48, p. 1815, 1977.

[15.17] H. Ryssel, H. Iberl, M. Blier, G. Prinkle, K. Habergerr and H. Kranz, "Arsenic Implanted Polysilicon Layers," *Appl. Phys.*, 24, p. 197, 1981.
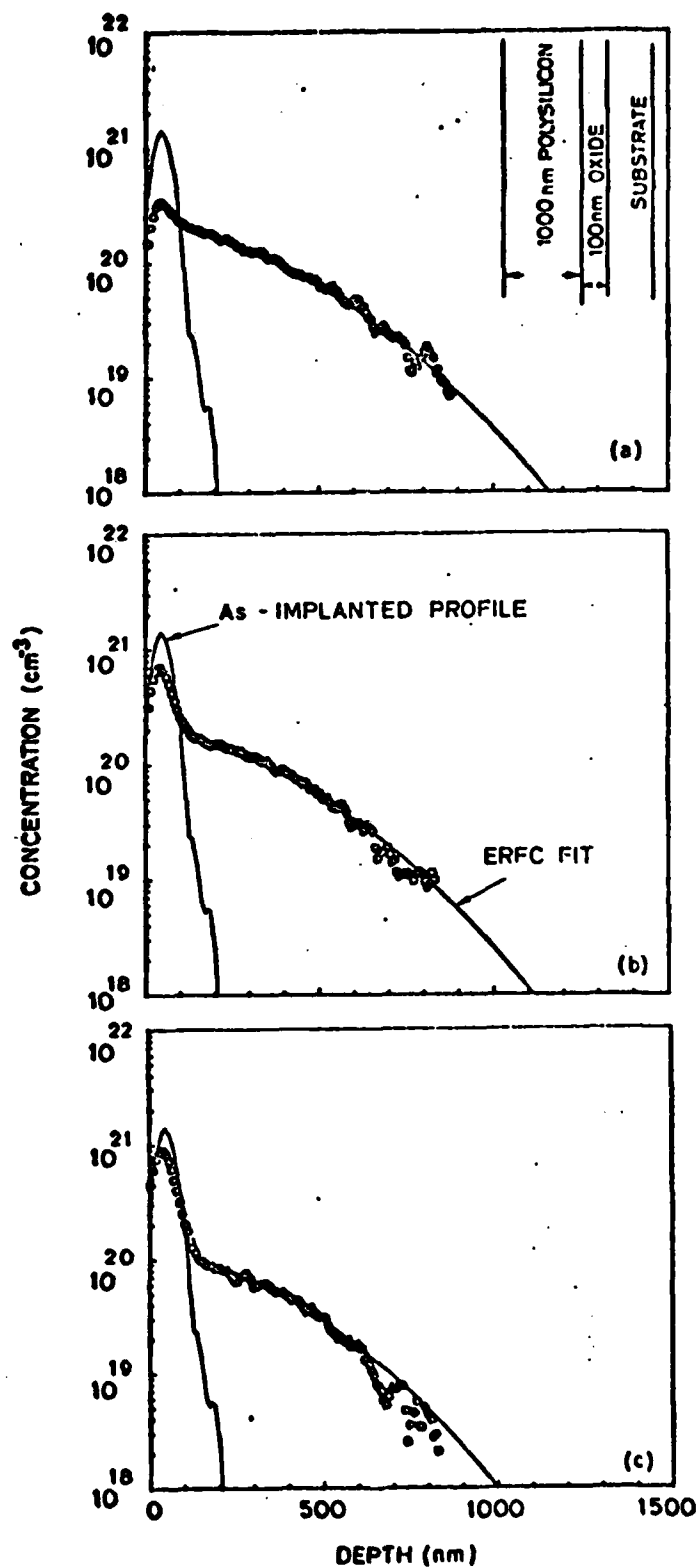
Fig. 15.1: Concentration profiles of arsenic in 1 μm polysilicon films of different grain sizes after annealing at 800°C for 6 hours in nitrogen. The initial profile of the arsenic introduced by a 40 KeV implant at a dose of 10^{16} cm^{-2} is also shown. The average grain sizes of the films are 0.21 μm (a), 0.36 (b) and 0.51 (c).
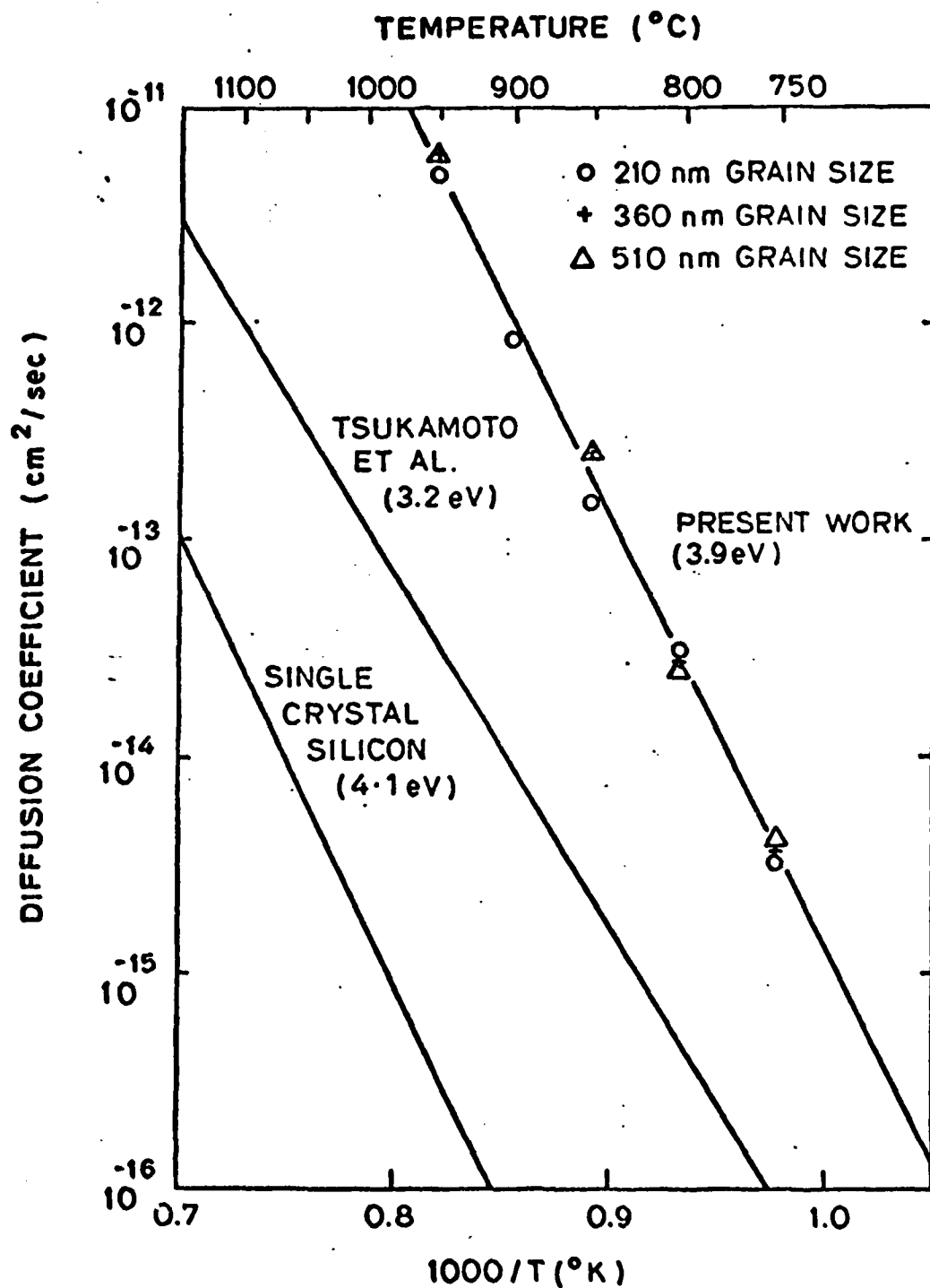
Fig. 15.2: The diffusion coefficient of arsenic in the polysilicon grain boundaries as a function of temperature. Also shown are the data of Tsukamot et al. [15.16] as well as the diffusivity of arsenic at low concentrations in single-crystal silicon [15.6].
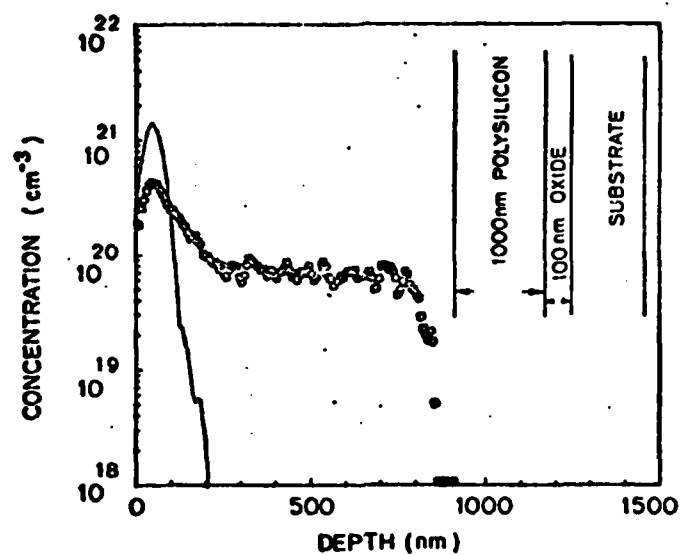
Fig. 15.3: Concentration profile of arsenic in a 1 μm polysilicon film of grain size 0.51 μm after annealing at 900°C for 1 hour in nitrogen. The broadening of the surface peak as compared to the initial profile allows an estimation of the value of the arsenic diffusivity in the interior of the grains.

Poly-Si ⟶

SiO₂ ⟶

Simulation region

(a)

arsenic

10^{16} at 40 KeV
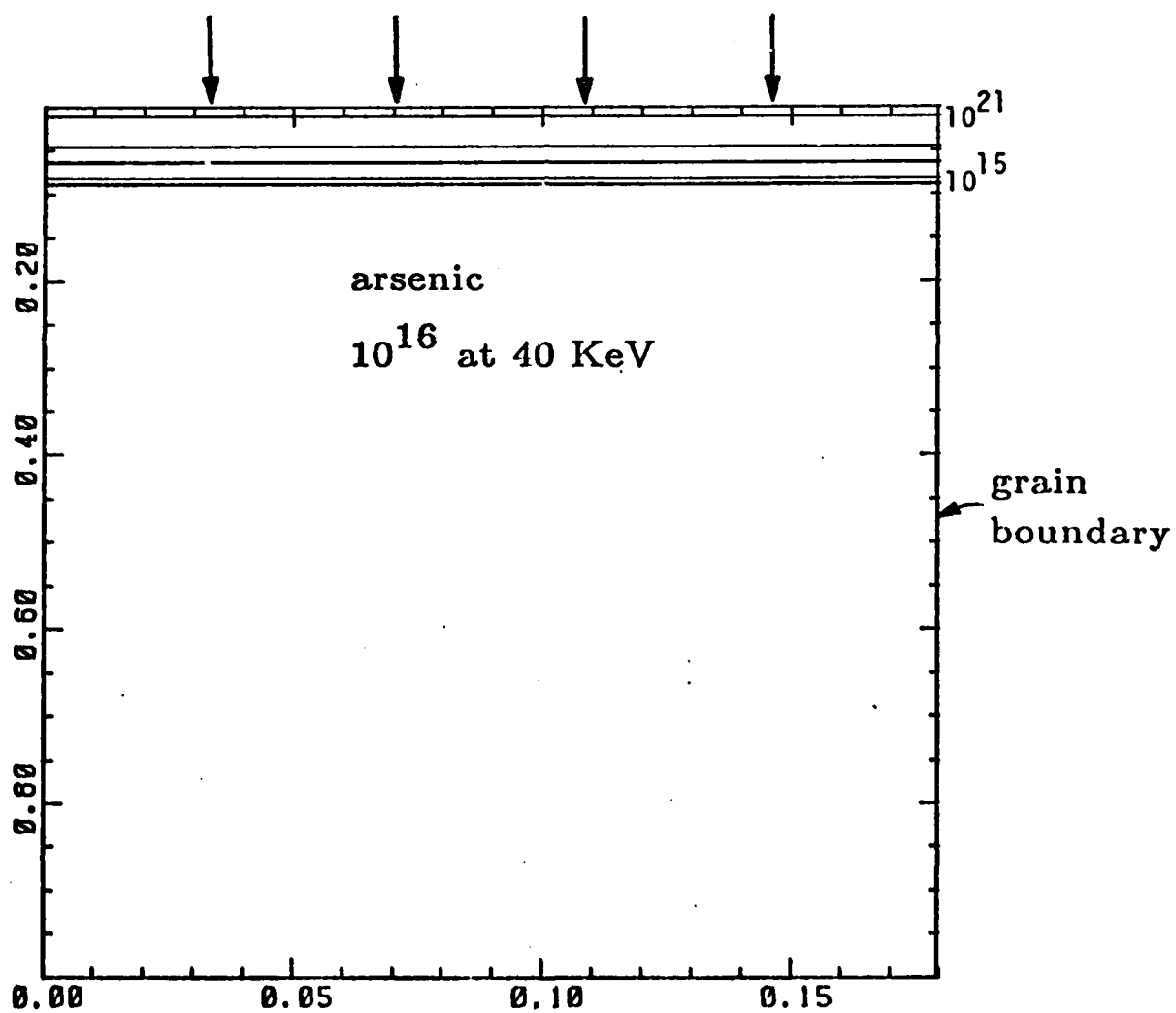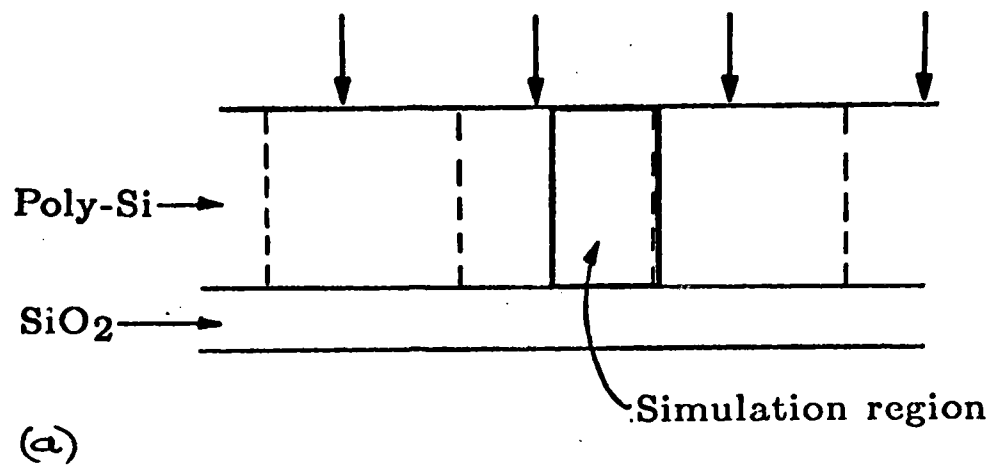
grain boundary

Fig. 15.4: Structure used in SUPRA simulations.

387

As, 1×10^{16} @ 40 keV

INERT AMBIENT ANNEAL
6 HOURS AT 800°C
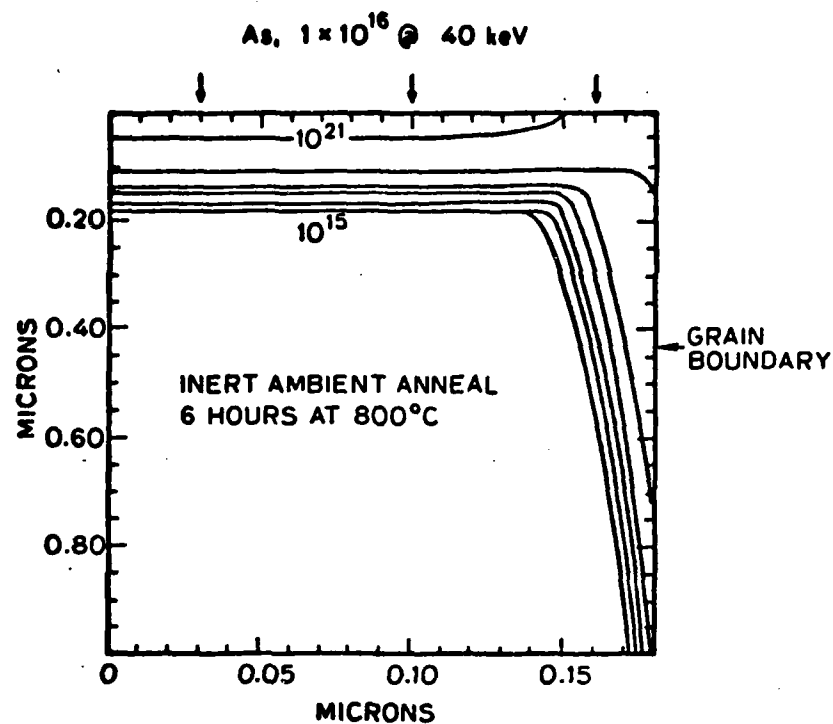
GRAIN BOUNDARY

Fig. 15.5: (a) Cross-sectional view of simulated arsenic contours in a single grain of polysilicon following implantation and diffusion. The grain boundary passes vertically along the right hand side.

□ RUTHERFORD BACKSCATTERING

— SIMULATED AVERAGE VERTICAL PROFILE
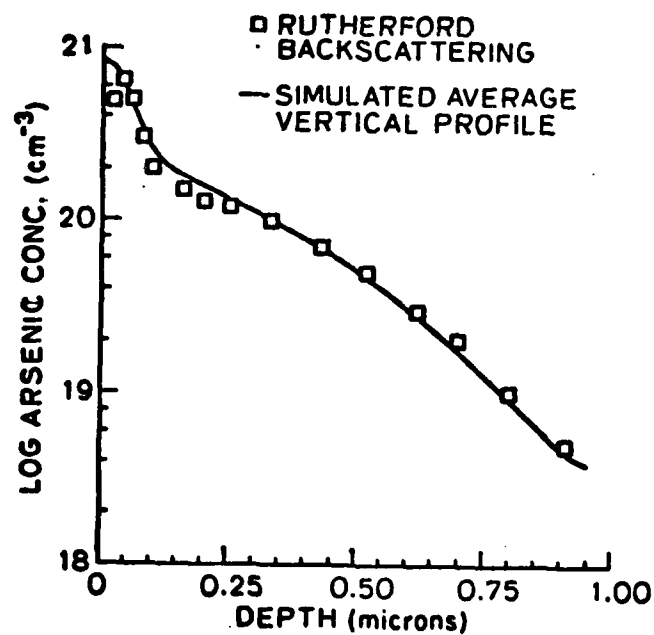
(b) Vertical profile of average arsenic concentration versus depth through the polysilicon layer.
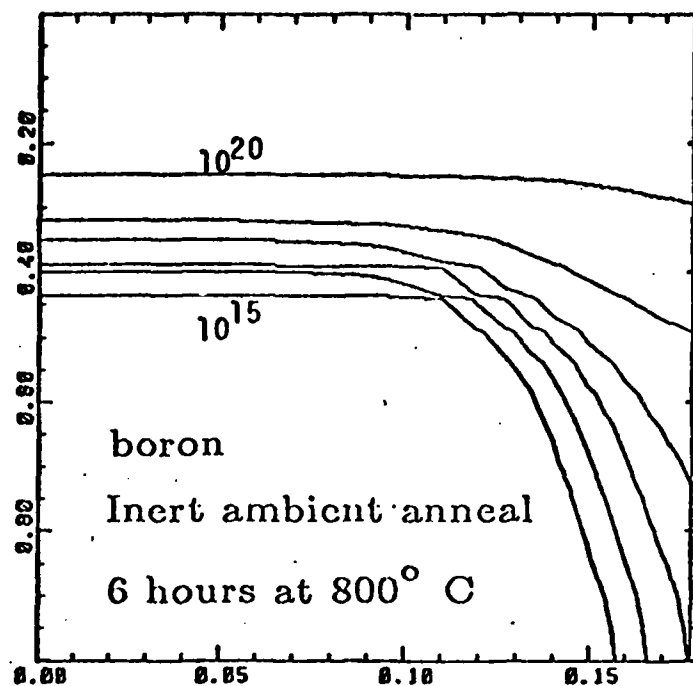
Fig. 15.6: (a) Simulated boron contours in a single grain of polysilicon after implantation and diffusion.
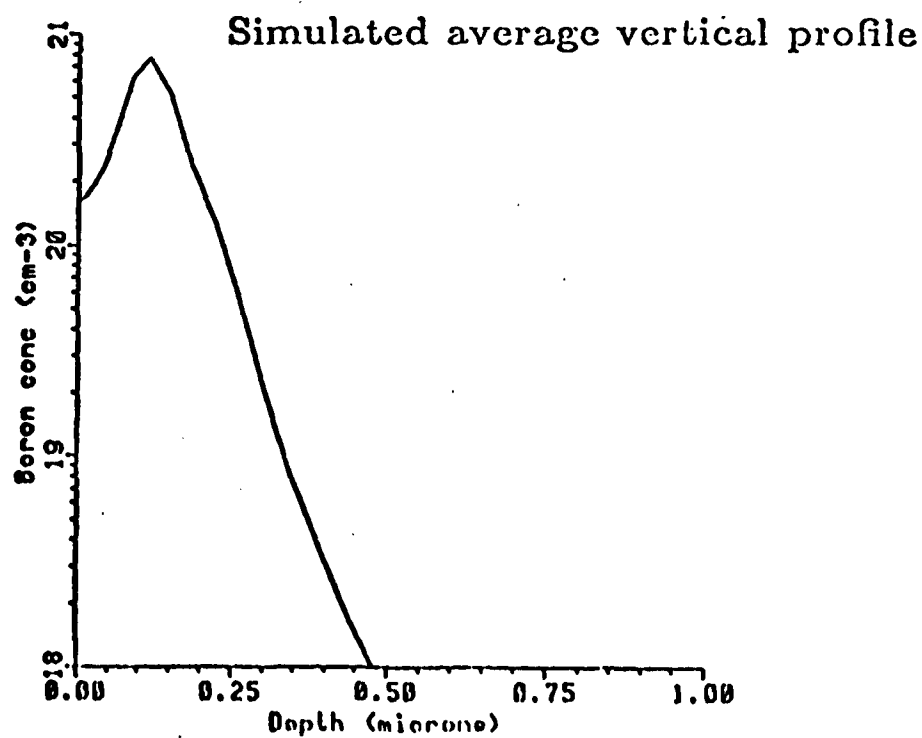
(b) Vertical profile of average boron concentration versus depth through the polysilicon layer.
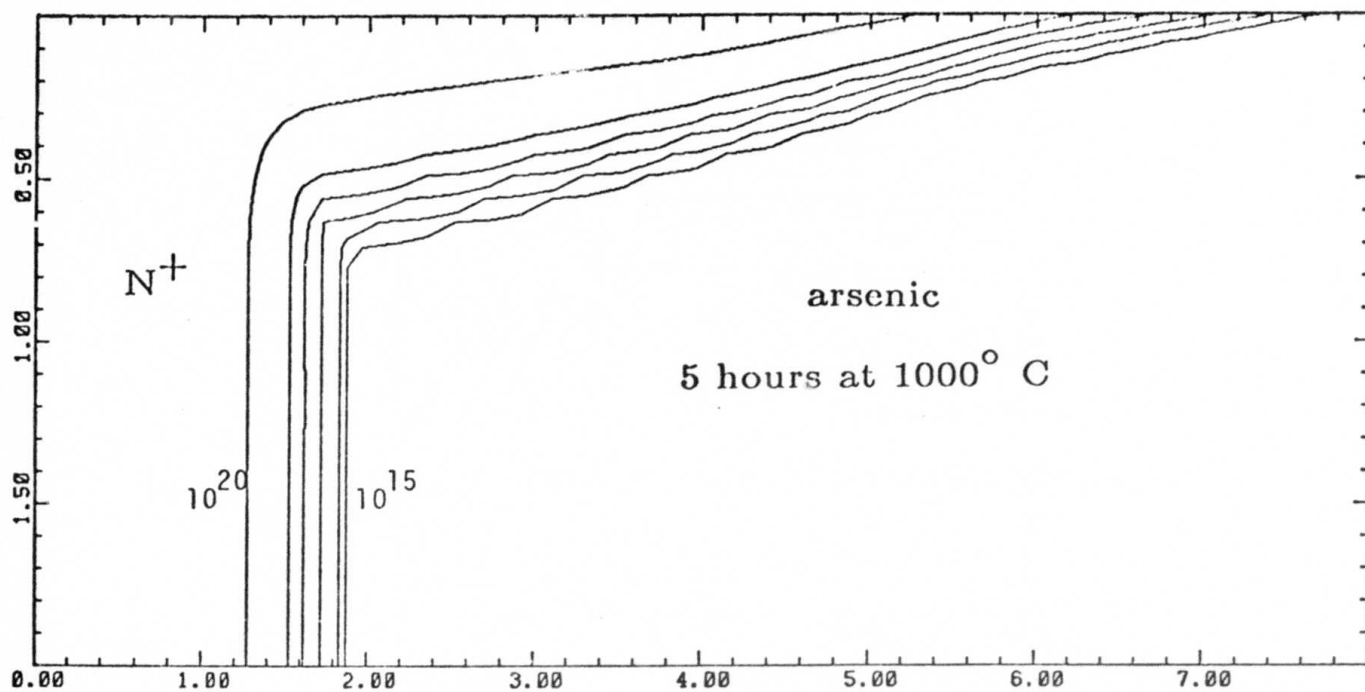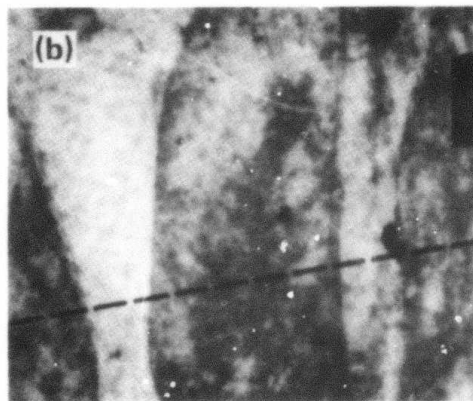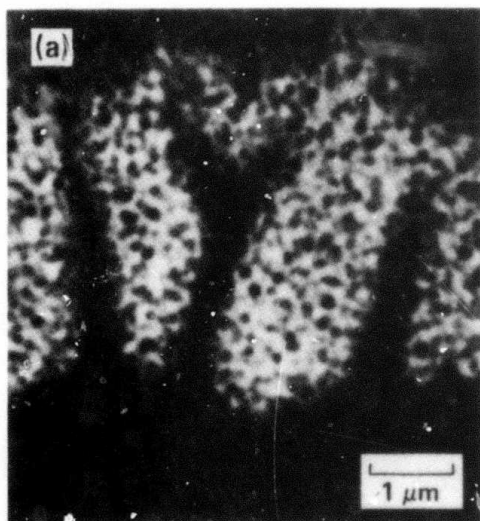
389

Fig. 15.7: Arsenic diffusion from a highly doped area on the left along a grain-boundary. Below are the experimental results of Johnson et al.

16.   PRACTICAL CONSEQUENCES OF DOPANT SEGREGATION

R. W. Barton, C. R. Helms, W. A. Tiller

16.1 SURVEY OF DOPANT SEGREGATION EFFECTS ON DEVICE BEHAVIOR

In previous periods we have experimentally and theoretically modeled the segregation of phosphorus and arsenic at the Si-SiO$_2$ interface. A version of the model has been incorporated into SUPREM III. In the following sections of this report some of the practical device consequences of the segregation effect are discussed and examples of calculated profiles using an analytic solution to the dopant segregation problem presented.

16.1.1 Effects on Interface Properties

In previous work we presented a few atomic models for the binding site of phosphorus in the Si/SiO$_2$ interface. These were based on idealized atom arrangements in a silicon <111> interface which, to first order, we assumed were not disrupted by the presence of phosphorus. Several second order effects due to phosphorus segregation, however, may account for some anomalous properties of the interface that have been reported in the literature.

We concluded in previous work that the phosphorus has no first order effect on electrical properties of the Si/SiO$_2$ interface. The dopant related effects reported by Snell [16.1] were considered too small in comparison to the total density of segregated phosphorus. To second order, however, Snell's observations would be consistent with the properties of a phosphorus rich layer in the inter-face.

Snell observed, for instance, that the density of interface traps will in-crease with phosphorus doping level. These defect states, with energies in the silicon band gap, have been associated with dangling bonds in the interface [16.2]. Since their density is so small, we did not associate them with a driving force

for segregation. Most of the phosphorus atoms, we concluded, retain their four-fold coordination in the interface (with bonds either to silicon or oxygen atoms). It remains true, however, that a dangling bond attached to a phosphorus atom will be more stable than one attached to silicon and so the probability of finding a dangling bond should increase in the presence of high P concentrations. The fraction of interface phosphorus atoms associated with interface electron traps appears to be about two percent.

The extra dangling bonds that appear in highly doped samples, however, might not be associated directly with phosphorus atoms. It is possible that the presence of large P concentrations strains the lattice in the interface region and leads to the formation of extra electron defects on silicon atoms.

The indirect effect of phosphorus segregation on the properties of interface silicon atoms might also be used to account for Ho's observations of dopant enhanced oxidation rate [16.3]. Ho observed that, in heavily doped samples, the linear oxidation rate of silicon will increase. The linear rate constant, in contrast to the parabolic rate onstant, is thought to be associated with a reaction mechanism at the Si/SiO_2 interface. In a detailed kinetic analysis, Ho reported that the dopant enhancement effect occurs through increases in the pre-exponential rate constant. The activaiton energy of the interface reaction is not significantly affected by doping level. This can be interpreted to mean that the concentration of reaction centers is increased in the heavily doped inter-face. If the phosphorus contributes to the number of dangling bonds in the interface, it might also be increasing the number of sites where oxygen can become attached. This was modeled by Ho as an increase in the silicon vacancy concentration.

Having identified the dopant segregation phenomenon in the Si/SiO_2 inter-face, it would not be unreasonable to expect that similar phenomena occur in the

other silicon interfaces. Evidence has already been found, for instance, for the segregation of arsenic to the grain boundaries of polycrystalline silicon [16.4]. This segregation has been found to significantly reduce the concentration of active charge carriers in a thin polycrystalline film. Evidence has also been found for the segregation of arsenic to the free surface of silicon [16.5]. (The evidence is sketchy, however, since the dopants evaporate rapidly from the crystal surface). High dopant concentrations will probably have an effect on the kinetic properties of this surface, influencing silicon epitaxial growth rates and the rates of dopant incorporation from gaseous sources.

16.1.2 Effects on Dopant Distributions in Silicon

The understanding and control of dopant concentration profiles in a silicon wafer is a necessary part of device fabrication. Many practical problems assoc-iated with these concentration profiles, however, have challenged our ability to understand fundamental diffusion processes. Many of the profiles can only be understood through the use of diffusion constants that depend variously on con-centrations, concentration gradients, or position in the sample (see reviews by Hu [16.6] or Tsai [16.7]). Chemical potential fields, related for instance to dopant segregation, also need to be considered in the diffusion models.

A typical SUPREM simulation is shown in Fig. 16.1. Two phosphorus profiles are shown here, one resulting from an ion implantation step, the other showing the effect of a subsequent, high-temperature, "drive-in" anneal. The program also generates from these profiles device parameters such as sheet resistance and junction depth.

The SUPREM program has now been revised to account for the effects of inter-face segregation. Fig. 16.1, however, does not show the phosphorus pile-up that we would expect to observe in the Si/SiO$_2$ interface. As a consequence, this simulation may be overestimating the amount of phosphorus that exists in the bulk of

393

the silicon. Also, if the segregated phosphorus is electrically inactive, the
simulation may be predicting sheet resistances that are too low in the silicon
wafer.

We have already observed significant changes in a silicon dopant profile due
to the effects of interface segregation. In Fig. 16.2, we show three Auger
sputter profiles which illustrate the evolution of dopant distributions near the
interface with oxidation time. Of particular interest is the profile obtained
after ten minutes of oxidation. We have noted that this sample contains more
phosphorus in the interface than can be accounted for by redistribution from the
SiO_2.

We proposed that the interface contains phosphorus which had diffused into
the interface from the bulk of the silicon. We attributed the gradients on the
silicon side of the interface to a diffusion profile which accompanies the flux
of phosphorus.

The revised model in *SUPREM III* will be able to simulate *diffusion profiles*
such as those shown in the ten-minute oxidation sample. However, many of the
expected effects can be predicted from an analytical solution to a simplified
version of this diffusion problem. This solution will be described in the next
section. The solution will then be compared to our Auger profiles in the follow-
ing section.

16.2 INTERFACE FIELD EFFECTS ON SOLUTE DISTRIBUTIONS DURING CRYSTAL GROWTH - THEORY

16.2.1 Introduction

Solute redistribution is an important problem in many practical studies of
phase transformations. It originates in the tendency of a growing phase to
either reject or absorb impurites as they are encountered in the parent solution.
It results in concentration gradients that can extend from the moving interface
far into either phase. Dopant gradients that are expected to arise from the

394

redistribution during oxidation are reproduced in Fig. 16.3. A quantitative understanding of these dopant distributions is important since they can influence electrical behavior in MOS devices. The redistribution problem is equally important in the semiconductor crystal growth process--where resulting dopant gradients influence the velocity and morphology of the crystal growth front, while simultaneously determining the doping level in the product crystal.

A quantitative analysis of the redistribution effect begins with Fick's laws of diffusion:

$$J = -D \frac{\partial C}{\partial x} \tag{16.1}$$

$$\frac{\partial C}{\partial t} = -\frac{\partial J}{\partial x} \tag{16.2}$$

where J represents the flux of a diffusing species, C is concentration, and D is the diffusion coefficient; x and t represent position and time, respectively.

These equations must be solved in the vicinity of a moving boundary where the concentrations in the two phases (1 and 2) are related by an equilibrium coefficient:

$$k = \frac{C_1}{C_2} \tag{16.3}$$

In order to solve equations 16.2 and 16.3 for C(x,t), however, it is easier to change to a frame of reference fixed in the boudary, where the solute atoms will appear to approach the origin with an additional flux given by boundary velocity v.

$$J = -D \frac{\partial C}{\partial x} - vc \tag{16.4}$$

$$\frac{\partial C}{\partial t} = \frac{\partial}{\partial x} \left(D \frac{\partial C}{\partial x} \right) + v \frac{\partial C}{\partial x} \tag{16.5}$$

These were the equations solved by Grove, Lestiko and Sah [16.9] for the dopant redistributions that occur during oxidation. They assumed that the velocity, v, had a parabolic time dependence.

$$v = at^{1/2} \tag{16.6}$$

Equations 16.4 and 16.5, however, do not take into account the chemical forces associated with surface segregation. As discussed previously, these forces, particularly the interface fields, can lead to non-uniform dopant distributions in an extended region near the interface. During a diffusion process, these forces can lead to additional solute fluxes that will drastically alter the expected solute redistribution profile. These extra fluxes are driven by the gradients in single particle potentials, $\mu^\phi(x)$, that are associated with the interface fields. The fluxes from the field will be proportional to an Einstein mobility: D/kT. The more complete diffusion equations, then, may be written as:

$$J = -D \frac{\partial C}{\partial x} - vc - \frac{DC}{kT} \frac{\partial}{\partial x} \left(\mu^\phi(x) \right) \tag{16.7}$$

$$\frac{\partial C}{\partial t} = \frac{\partial}{\partial x} \left(D \frac{\partial C}{\partial x} \right) + \frac{\partial C}{\partial x} \left(v + \frac{D}{kT} \frac{\partial \mu^\phi}{\partial x} \right) + \frac{DC}{kT} \frac{\partial^2}{\partial x^2} \left(\mu^\phi(x) \right) \tag{16.8}$$

Tiller and Ahn [16.10] have shown how the existence of interface fields can alter the concentrations of dopant incorporated into a growing crystal. They solved equations 16.3 and 16.7 for a crystal growing at constant velocity (the steady state solution). They used a simple form for interface fields.

$$\mu^\phi(x) = \beta \exp (\alpha x) \tag{16.9}$$

Their resulting concentration profile is very dependent on growth velocity. At low velocities the dopant is able to diffuse relatively quickly, and a close approximation to an equlibrium dopant profile is maintained near the interface.

396

At higher velocities, however, solute fluxes become dominated by the motion of the interface and equilibrium distributions are not achieved. The result can be measured as an effective distribution coefficient, k', that varies with growth velocity and crystal orientation. These effects have been observed by Hall [16.11]. A comparison between Hall's measurements of the effective distribution coefficient and those predicted by Tiller and Ahn is shown in Fig. 16.4.

Using parameters defined in equation 16.9, Tiller and Ahn showed that the interface distribution will approximate an equilibrium distribution during growth as long as

$$V << \frac{\beta \alpha D}{k_B T} \tag{16.10}$$

This condition is in fact satisfied during the silicon oxidation process.* As a result, we do not expect velocity dependent effects, similar to those shown in Fig. 16.4 to occur during oxidation.

Additional interface field effects can occur, however, during the initial oxidation process. In this initial transient period, while the solute is first beginning to pile-up in front of the moving interface, we will find that an interface field can cause solute to move in directions that are just the opposite of those expected from normal redistribution theory. These additional fluxes will be important in our understanding of the ten-minute oxidation profile shown in Fig. 16.2.

16.2.2 A Segregation Model for the Initial Transient

The intitial transient solute profile can, in principle, be obtained from a solution to equation 16.8. No exact solution, however, has yet been given for this

---

* $V \simeq 20$ A/min. $\qquad$ $D \simeq 5 \times 10^{13}$ A^2/min.

$\quad$ B $\simeq 0.3$ eV $\qquad\qquad$ kT $\simeq 0.1$ eV

$\quad$ $\alpha \simeq -.05$ A^{-1} $\qquad$ 397

equation. We will instead demonstrate the principle effects of segregation on the initial transient with a few simple approximations.

Our approximations center on a simplified functional form for the interface field. Instead of Tiller and Ahn's exponential form, we will employ a square well chemical potential, illustrated in Fig. 16.5. In equilibrium, this chemical potential should result, as shown, in a box-like distribution of dopant near the interface. The object of our diffusion calculations will be to determine how fast this box can be filled with phosphorus from an initially flat concentration profile.

The step-like chemical potential of Fig. 16.5 allows us to treat the interface as if it were a separate chemical phase with a greater affinity for phosphorus. Since the chemical potentials are uniform in each phase, the much simpler form for the diffusion equation can now be used to solve for the concentration gradients (equation 16.5 instead of 16.8). The segregation phenomenon will manifest itself through a boundary condition: a distribution coefficient, like eqn. 16.3, will apportion the dopant between the interface plane and the bulk plane.

The square well chemical potential is a severe approximation to the sort of interface fields expected here. However, if the width is chosen to correspond to a single lattice distance, the square well may be a very good approximation for the monolayer model of segregation. We have chosen to leave the width, δ, as an adjustable parameter. The total volume of the equilibrium distribution, given by the product $k\delta$, is a material parameter that can be determined from Auger sputter profiles.

In order to simplify the problem further, we have made three more assumptions: (a) We assume that the SiO_2 is impermeable to phosphorus (the actual distribution coefficient between silicon and SiO_2 is greater than ten.), (b) We assume that the growth velocity is constant (as it usually is in the early stages

of silicon oxidation), and (c) we assume that the interface region is so narrow that concentration gradients are insignificant in that region. We are assuming, in other words, that the diffusion constant in the interface phase, D^I, is infinite. These assumptions allow us to concentrate our efforts on the determination of dopant profiles in the bulk of the silicon.

We will now describe our boundary conditions for the diffusion process in silicon. The initial condition should appear as in Fig. 16.6. The concentration is uniform at every point--except for a finite discontinuity in the interface. At time t = 0+, therefore, no phosphorus has diffused into the boundary. As in many diffusion problems, the calculated flux has a pole at time zero.

Our expression for the initial condition is:

$$c^I (t = 0) = C_\infty \qquad c^{Si} (t = 0) = \begin{array}{l} C_\infty/k \quad \text{at } x = 0 \\ C_\infty \quad \text{at } x \neq 0 \end{array} \qquad (16.11)$$

We expect the concentration of solute in the interface phase to increase uniformly as solute is dumped into it from the silicon (see Fig. 16.7). These changes will be governed by the solute conservation principle:

$$\frac{\partial}{\partial t} (c^I) = - \frac{J^{Si}(x = 0)}{\delta} \qquad (16.12)$$

where δ is the width of the interface layer. If we note in addition the proportionality between C_I and C_{Si} we can write a boundary condition at x = 0 that employs only silicon concentration variables:

$$\frac{\partial}{\partial t} [C^{Si}] = \frac{1}{k\delta} [d \frac{\partial C^{Si}}{\partial x} + vC^{Si}] \qquad (16.13)$$

where $k = C^I/C^{Si}(x = 0)$.

The other boundary condition is determined by the uniformity that should exist in the silicon at distances far away from the interface:

$$c^{Si} (x = \infty) = \text{constant} \qquad (16.14)$$

The solution specified by equations 16.5, 11, 13, and 14 appears as follows:

$$C(x,t) = C_\infty + C_\infty \left\{ \; a\,(-\rho^{1/2})\; \text{erfc}\left[\frac{\sigma^{1/2}}{2t^{1/2}} + (\rho t)^{1/2}\right] \qquad (16.15) \right.$$

$$+ b\,(-2\beta-\rho^{1/2})\; \exp\left[(2\beta)\sigma^{1/2} + (4\beta^2 + 4\beta\rho^{1/2})t\right]$$

$$\times\; \text{erfc}\left[\frac{\sigma^{1/2}}{2t^{1/2}} + (2\beta + \rho^{1/2})\, t^{1/2}\right]$$

$$+ c\,(\rho^{1/2})\; \exp\left[-2(\rho\sigma)^{1/2}\right]\; \text{erfc}\left[\frac{\sigma^{1/2}}{2t^{1/2}} - (\rho t)^{1/2}\right]$$

$$\left. + d\,(1-(\rho\sigma)^{1/2}+ 2\rho t)\;\; \exp\left[-2(\rho\sigma)^{1/2}\right]\; \text{erfc}\left[\frac{\sigma^{1/2}}{2t^{1/2}} - (\rho t)^{1/2}\right] \right\}$$

. . . .

where

$$a = \frac{1}{2\rho^{1/2}} \qquad\qquad\qquad\qquad\qquad c = -a - b \qquad (16.16)$$

$$b = \frac{-q}{2\beta + 2\rho^{1/2}} - \frac{\rho^{1/2}}{2(\beta + \rho^{1/2})^2} \qquad\qquad d = \frac{\beta}{\beta + \rho^{1/2}}$$

16.2.3 Results

In Figs. 16.8 and 16.9, we illustrate the calculated solution to the initial transient problem (equation 16.15). In these figures, we have used a chemical potential well that would result in an equilibrium distribution that is 40 A wide and 12 times more concentrated in the interface than in the bulk silicon. The other parameters of interface velocity and phosphorus diffusion coefficient

correspond to published values for the oxidation of silicon at 900°C. (C_∞ = 1.2×10^{20} cm^{-3}, D = 4.25×10^3 Å/min, v = 3 Å, from references 16.12 and 16.13).

The first frame (Fig. 16.8a) shows a phosphorus distribution after .005 minutes of oxidation; it closely corresponds to our initial condition. The second frame (16.8b) shows the interface beginning to fill up with phosphorus, accompanied by steep concentration gradients in the silicon. In Fig. 16.8c, the near interface region has nearly reached equilibrium, although a small concentration gradient persists deeper in the silicon wafer.

Fig. 16.9 illustrates the final stages of the initial transient, after 50 and 200 minutes of oxidation, respectively. Note that in the time between these two profiles, the gradient near the interface has changed from positive to negative values. This transition is coincident with a reversal in the direction of phosphorus flux in the region of the interface. As the chemical potential well fills up with phosphorus, the normal redistribution process that accompanies the rejection of dopant from the SiO$_2$ begins to dominate the profiles. The redistribution does not begin to cause a phosphorus pile-up in the bulk of the silicon, however, until after at least 250 Å of oxide have been grown.

16.2.4 Conclusions

One of the objects of our study has been to determine the rate at which interface concentrations increase during the early stages of oxidation. This rate is given by the flux of solute from the silicon into the interface, and will be influenced by the velocity of the interface as well as by concentration gradients in the silicon. In the early stages of oxidation, however, the flux is dominated by concentration gradients. This has been shown by comparing profiles such as in Fig. 16.8 to profiles obtained by setting the interface velocity to zero. After one minute of oxidation the phosphorus distributions are essentially identical. After 20 minutes of oxidation (at 900°C) the stationary interface has only

twenty percent less phosphorus than the moving interface. It is only in the later stages of oxidation, when the concentration gradients are small, that the movement of the interface has a predominant effect on the evolution of the inter- face concentration. All of this is in contrast to the former theories of redis- tribution, where the dopant profile was directly related to interface movement.

We have also studied how the shape of the potential well affects the evolu- tion of the dopant profile. Since it is difficult to measure interface width directly in the laboratory, it was felt that the theoretical width, δ, could be determined from comparisons between theoretical and experimental dopant profiles in silicon. Unfortunately, we have found that, as long as the volume of the potential well remains constant, the relative height and width of the well has no significant effect on the evolution of the profiles in the silicon. Presumably, this is because the impurity fluxes are controlled by concentration gradients that extend far into the bulk of the solid, where they are not affected by the interface field. For this reason, as long as one is not concerned with the exact distribution of solute in the interface region, the square well may accurately model the effects of many different interface segregation phenomena.

16.3 DOPANT SEGREGATION EFFECTS ON THE SOLUTE DISTRIBUTION DURING OXIDATION - PREDICTIONS AND APPLICATIONS

16.3.1 Comparisons with Auger Sputter Profiles

The initial transient study was first motivated by the need to understand the Auger sputter profiles obtained after 10 minutes of oxidation at 900°C (see Fig. 16.2). The phosphorus pile-ups observed in these samples contain an unusu- ally large amount of phosphorus and are accompanied by a depletion in phosphorus concentrations in the bulk of the silicon. One of the profiles is repeated in Fig. 16.10 along with the calculated results of our diffusion analysis.

Despite the crude "square well" approximation employed in our diffusion cal- culation, fairly good agreement is found between the theoretical and experimental

402

profiles shown in Fig. 16.10. The theory predicts approximately the total amount of phosphorus that has collected in the interface region, and also shows the concentration gradient in the depletion region. This prediction is based on the same values (taken from the literature) of interface velocity and diffusion coefficient as those used in Fig. 16.8. The volume of the potential well has also been determined by measuring an equilibrium phosphorus distribution in a sample that had been annealed in N_2 at 900°C.

By studying the evolution of the distibution theory with time, we have determined that the interface region will fill to within 90% of its equilibrium interface concentration after 90 minutes of oxidation at 900°C, and after 420 minutes of oxidation at 800°C. Both of these results have been confirmed by our Auger measurements.

Unfortunately, these changes in the dopant pile-up are not sensitive to the relative height and width of the chemical potential well. Therefore, we have n- been able to use the theoretical width, δ, as an adjustable parameter in our comparison. The actual width of the interface remains unknown since our experimental profiles are broadened by the effects of ion knock-on mixing. In order to show that a narrower and taller equilibrium distribution can be used to make the same predictions, we have analytically convoluted a theoretical profile and compared it also to the ten minute Auger profile (Fig. 16.11).

The theoretical curve in this figure is based on a potential well that has the same volume as that used in the calculation of Fig. 16.10, but is only 3 A wide. This profile was broadened with a 14 A wide gaussian distribution. Although the experimental profile does not show the same symmetry as our theoretical prediction, the two curves still have about the same volume, and both show concentration gradients in the silicon. A similar comparison could be made with a wider theoretical interface and a correspondingly narrow distribution.

16.3.2 Comparisons with Previous Redistribution Theories

The practical consequences of our segregation model can be evaluated by comparing its predictions to previous theories of dopant redistribution. We have already seen that segregation can give rise to depletions in dopant concentration on the silicon side of the interface. So far, however, we have only illustrated dopant profiles is heavily doped samples, and in regions very close to the interface. More extended dopant distributions are shown in Fig. 16.12 and 16.13.

The segregation profiles in these figures are compared to predictions from the SUPREM numerical simulation program (see previous section). The depth scales in these figures correspond more closely to distances important in device fabrication. In each figure, the previous redistribution theory predicts that phosphorus ought to be piled-up in front of the moving interface. Because of the growing segregation layer, however, our model predicts significant depletions in interface phosphorus concentrations. The concentration differences are typically about twenty percent, and can extend over distances of 1000 Å. Both figures show the results of a 900°C, 50 minute oxidation process, which result in oxide thicknesses of approximately 300 Å. Fig. 16.13, however, is based on a sample doped in the bulk to only 10^{17} cm^{-3} of phosphorus. At these lower dopings the phosphorus diffusion constant is smaller [16.12] and so the resulting concentration gradients are steeper and confined more closely to the interface. After 50 minutes of oxidation, the segregation layer is closest to its equilibrium concentration in the more heavily doped sample.

16.3.3 Effects on Implanted Layers

In the previous comparisons, we considered only changes in dopant profiles from silicon wafers that were assumed to be uniformly doped. Where the initial distribution is not uniform, however, or where there is a limited amount of impurity in the sample, the segregation effects can be much more significant. This

is illustrated for an ion implantation process in Fig. 16.14. This figure shows an estimate of the differences between implanted and diffused layers for models that either include or do not include segregation effects. Unfortunately, an exact prediction is not possible since our initial transient model is based on a uniform initial condition. The segregation effects included in SUPREM III will provide a more detailed comparison than is possible with our transient model.

In Fig. 16.4 we have used the phenomenological relations to estimate the amount of dopant that will become trapped in the interface after a sufficiently long anneal. This amount of segregated dopant was then subtracted from the total dose in the implanted layer. Notice that, in contrast to the profiles of Figs. 16.12 and 16.13, the ten percent difference in phosphorus concentrations extend throughout the diffused layer. This will be a reasonable prediction as long as the peak in the implanted profile was originally close to the interface.

Using the above estimate, SUPREM calculates that the profile with less phosphorus is also associated with a ten percent greater sheet resistance in the silicon (ρ = 21.3 ohms/square). The depth to the p-n junction is also 500 A less in the profile that accounts for interface segregation. Both of these effects can significantly influence device behavior.

REFERENCES

[16.1] J. Snell, Solid State Electron., 24, 135 (1981).

[16.2] R. B. Laughlin, J. D. Joannopoulos, and D. J. Chadi, in "The Physics of SiO$_2$ and its Interfaces," S. T. Pantelides, ed. (New York: Pergamon, 1978).

[16.3] C. P. Ho, J. D. Plummer, J. D. Meindl, and B. E. Deal, J. Electrochem Soc., 125, 665 (1978).

[16.4] M. M. Mandurah, K. C. Saraswat, C. R. Helms, and T. I. Kamins, J. Appl. Phys., 51, 5755 (1980).

[16.5] Stanley Perino, Stanford University, Private Communication.

[16.6] S. M. Hu, _Phys. Rev._, 180, 773 (1969).

[16.7] R. B. Fair, in _Semiconductor Silicon '77_, H. R. Huff, ed., (Princeton, NJ: The Electrochemical Society, 1977).

[16.8] D. A. Antoniadis, S. E. Hansen, and R. W. Dutton, "SUPREM II -- A Program for IC Process Modeling and Simulation," Stanford University, Technical Report No. 5019-2 (1978).

[16.9] A. S. Grove, O. Leistiko, Jr., and C. T. Sah, _J. Appl. Phys._, 35, 2695 (1964).

[16.10] W. A. Tiller and K. S. Ahn, _J. Crystal Growth_, 49, 483 (1980).

[16.11] R. N. Hall, _Phys. Rev._, 88, 139 (1952).

[16.12] R. B. Fair and J. C. C. Tsai, _J. Electrochem. Soc._, 124, 1107 (1977).

[16.13] C. P. Ho, J. D. Plummer, J. D. Meindl, and B. E. Deal, _J. Electrochem. Soc._, 125, 665 (1978).

PHOSPHORUS IMPLANT AND DIFFUSION
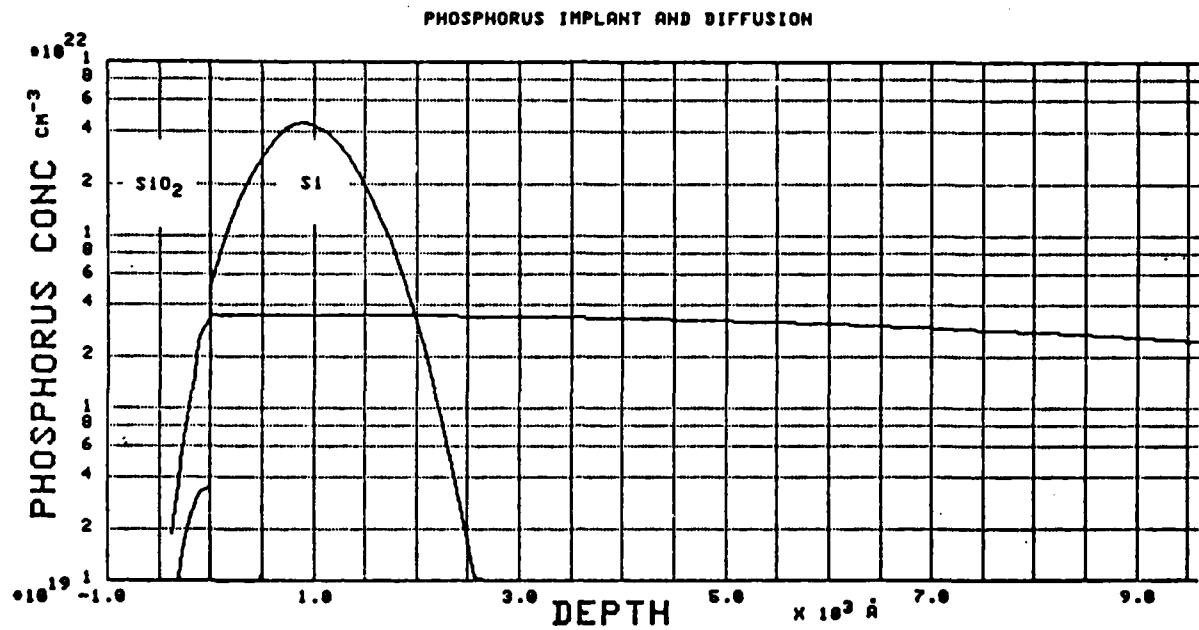
Fig. 16.1:   SUPREM II simulation, showing the results of phoshporus implantation (dose - 5 x 10^{16}, ion energy = 100 KeV) and a subsequent drive-in anneal (40 min., 1150°C, N$_2$).
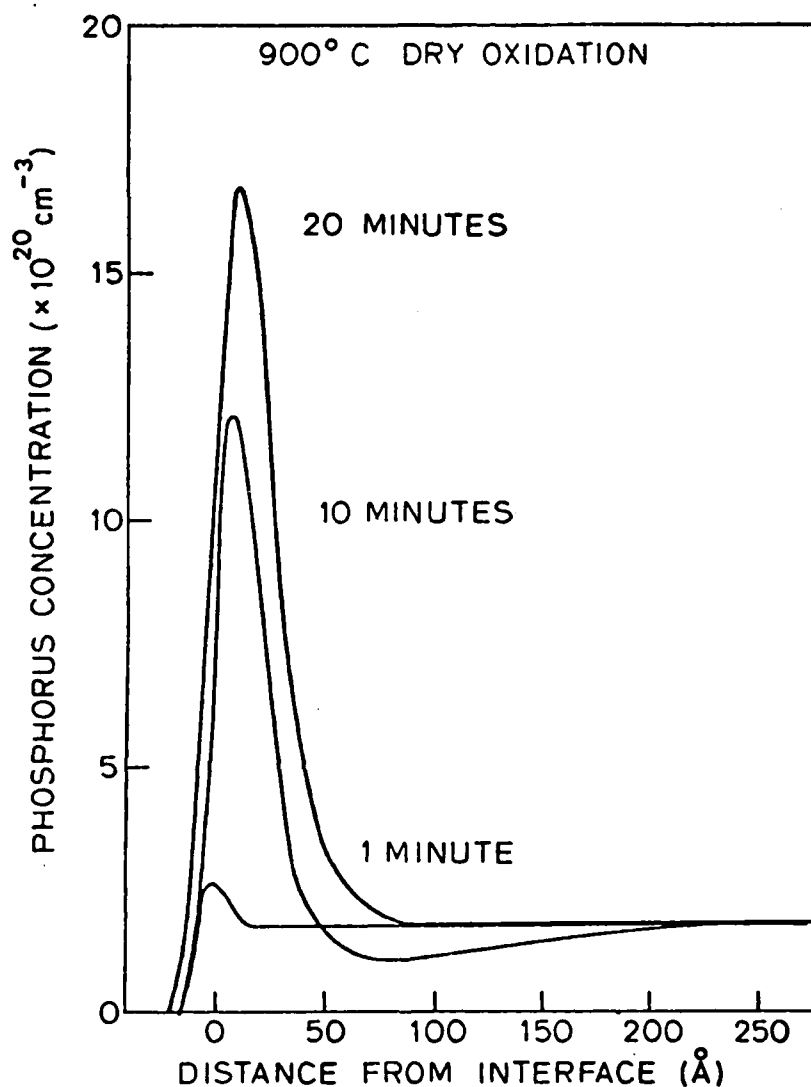
Fig. 16.2: Evolution of the phosphorus pile-up in the early stages of oxidation. We propose that much of the phosphorus in the interface after ten minutes of oxidation has reached there after diffusing out of the bulk silicon. Phosphorus concentrations in this sample are depleted on the silicon side of the interface.
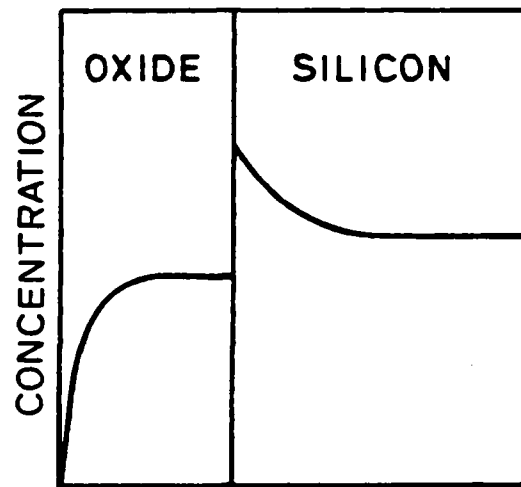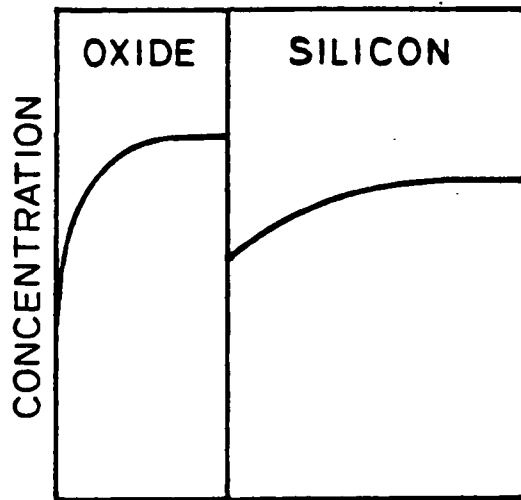
Fig. 16.3: Redistribution of impurities due to the thermal oxidation of silicon, k is the equilibrium distribution coefficient (c_{SiO_2}/c_{Si}).
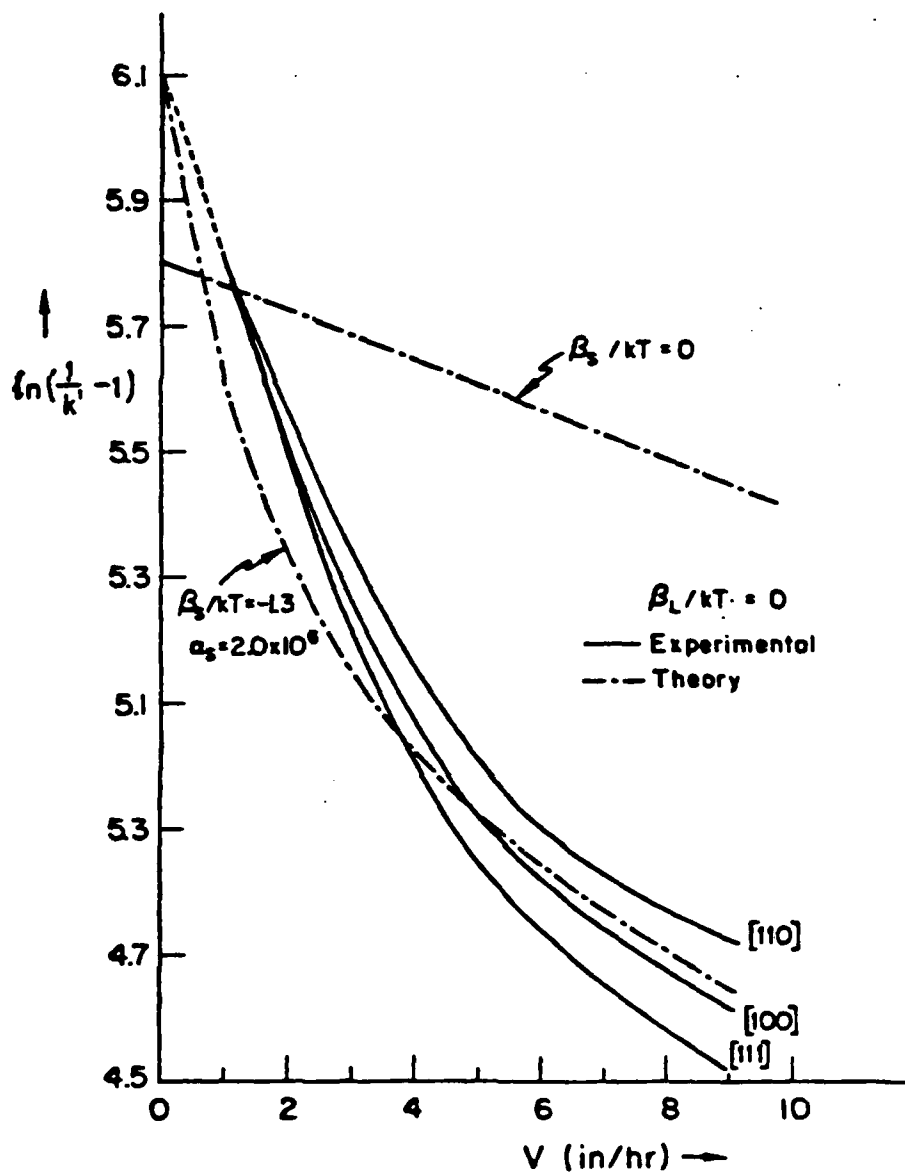
Fig. 16.4:   The effect of interface velocity on the effective interface
              distribution coefficient (k'). The solid lines show Hall's [16.3]
              experimental data; the dotted lines are the result of Tiller and
              Ahn's [16.4] diffusion theory. $\beta_S = 0$ implies the absence of an
              interface field.

Phosphorus Chemical Potential

$\mu^\phi_{(x)}$

X

Initial, Final Concentrations

C_0

$$C_0 = C_B \exp\left({-\Delta\mu^\phi}/{kT}\right)$$

C_B

X
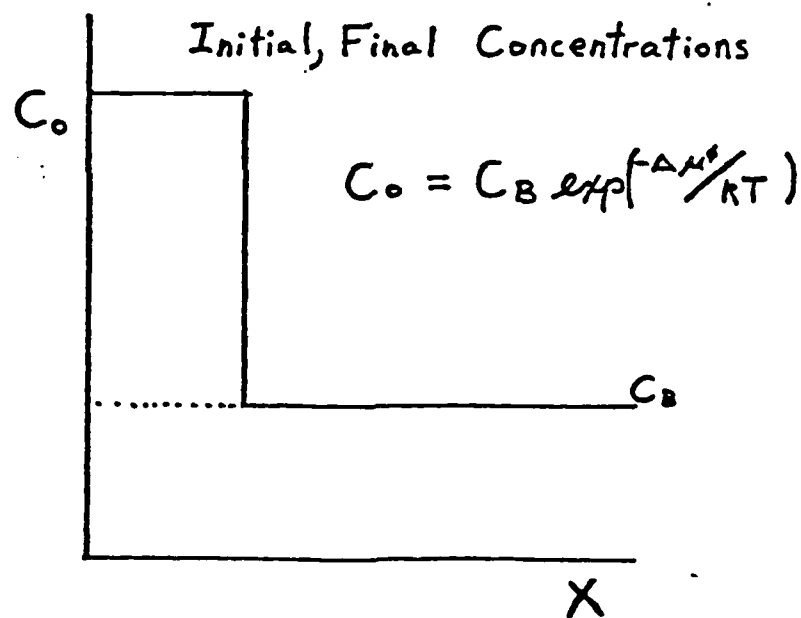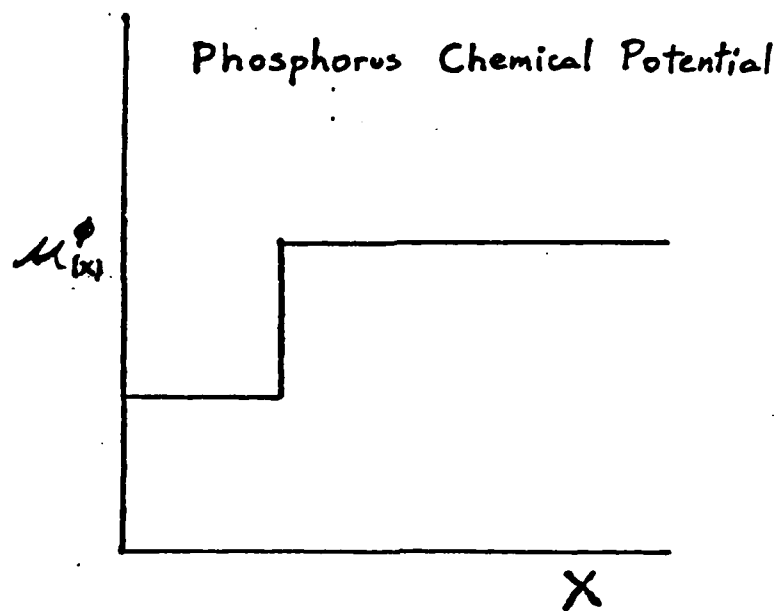
Fig. 16.5: The square well chemical potential, $\mu^\phi(x)$ used in our initial transient diffusion theory. (b) The uniform initial condition is shown as a dotted line; the final or equilibrium distribution is shaped like a box, the height of which is determined by the Boltzman equation.

411
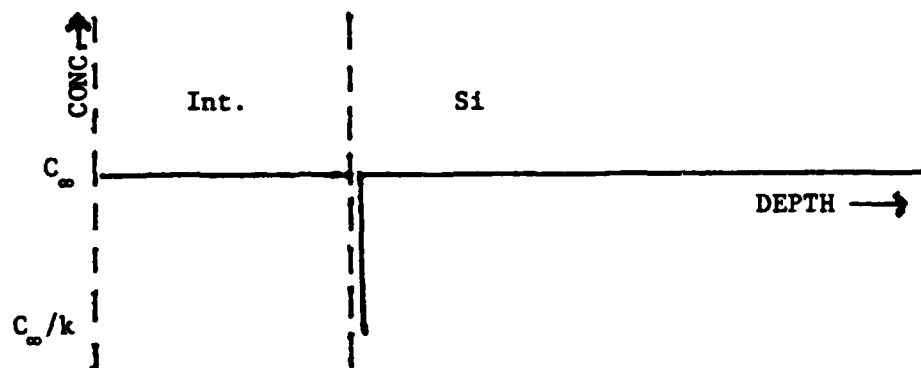
Fig. 16.6: The initial condition.
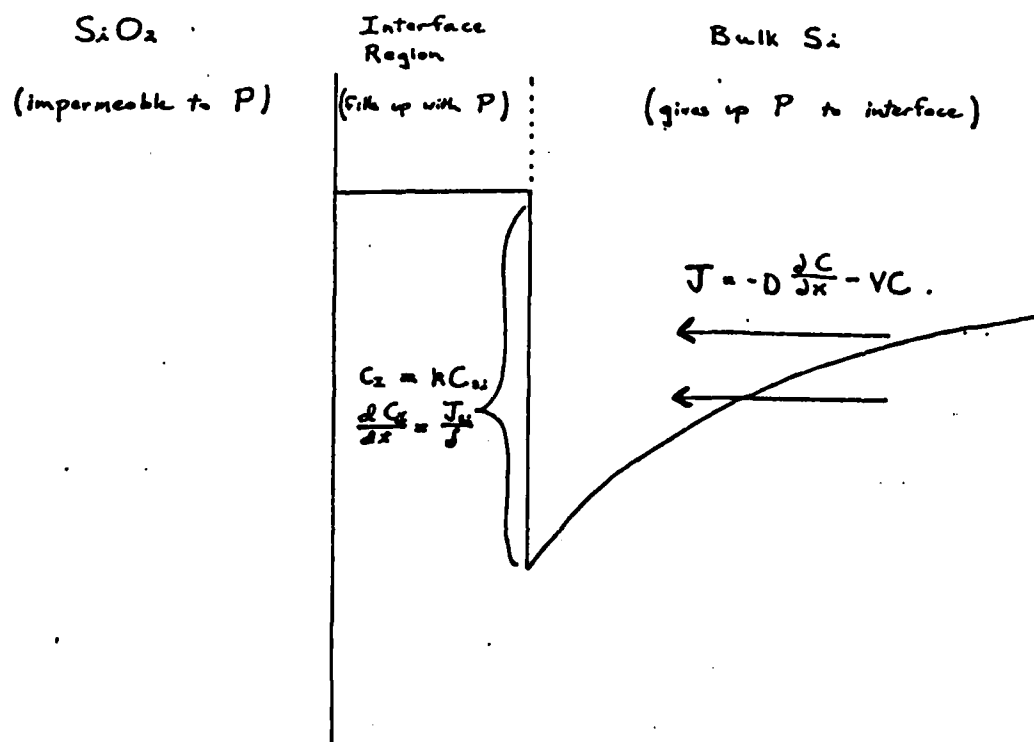
Fig. 16.7: Schematic of the diffused problem solved in our initial transient theory.

900° DRY OXIDATION
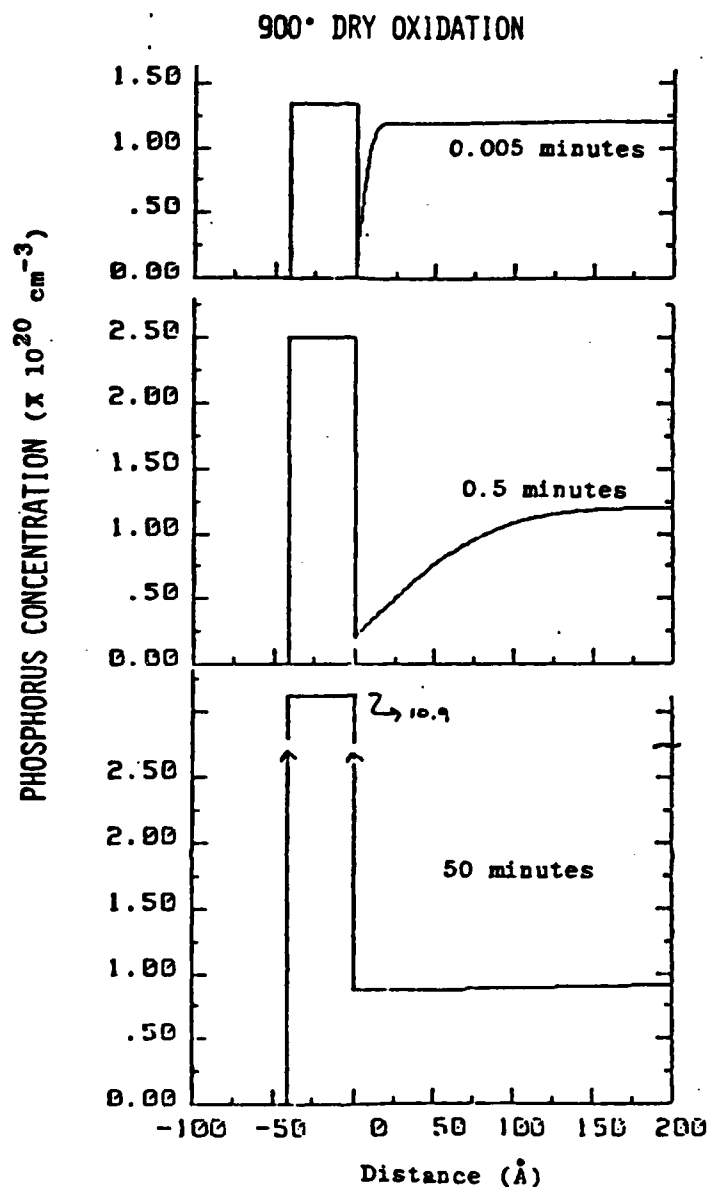
Fig. 16.8: The results of our initial transient theory in the early stages of
oxidation. These distributions shown the progression from our
initial condition to what is nearly an equilibrium distribution. δ
= 40 Å, k · δ = 480 Å.

900° DRY OXIDATION

PHOSPHORUS CONCENTRATION ($\times 10^{20}$ cm^{-3})

50 minutes

200 minutes

DISTANCE (Å)
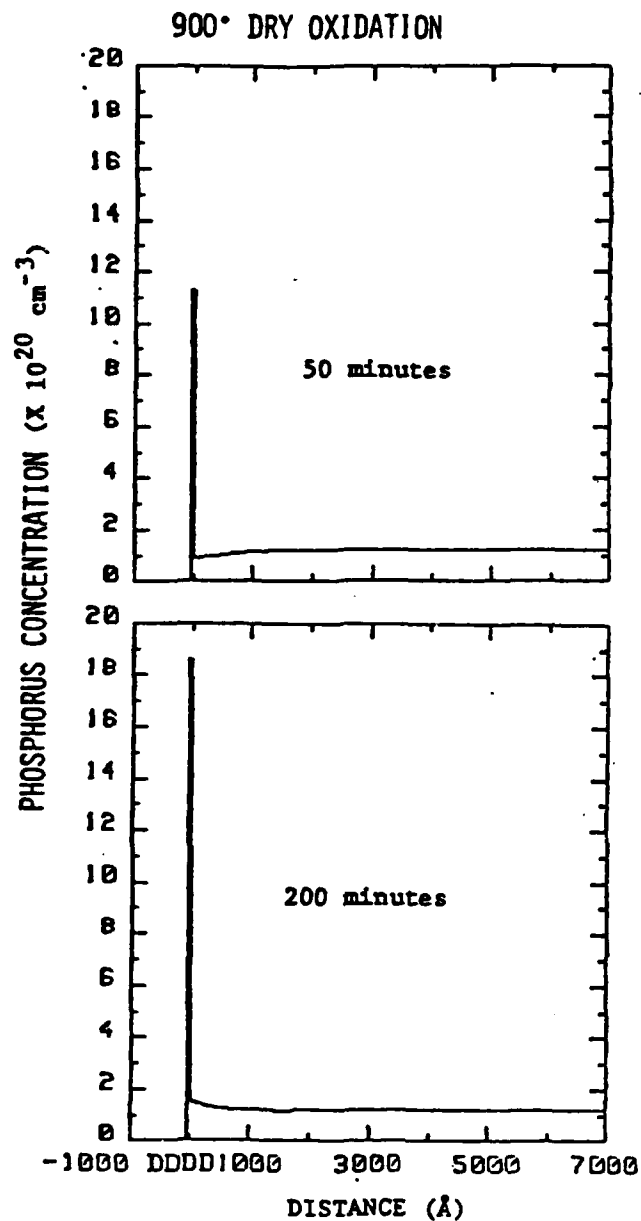
Fig. 16.9: The results of our initial transient theory in the intermediate stages of oxidation. Between 50 and 200 minutes the depletion region in the silicon becomes filled with phosphorus that has been rejected from the SiO$_2$.

414
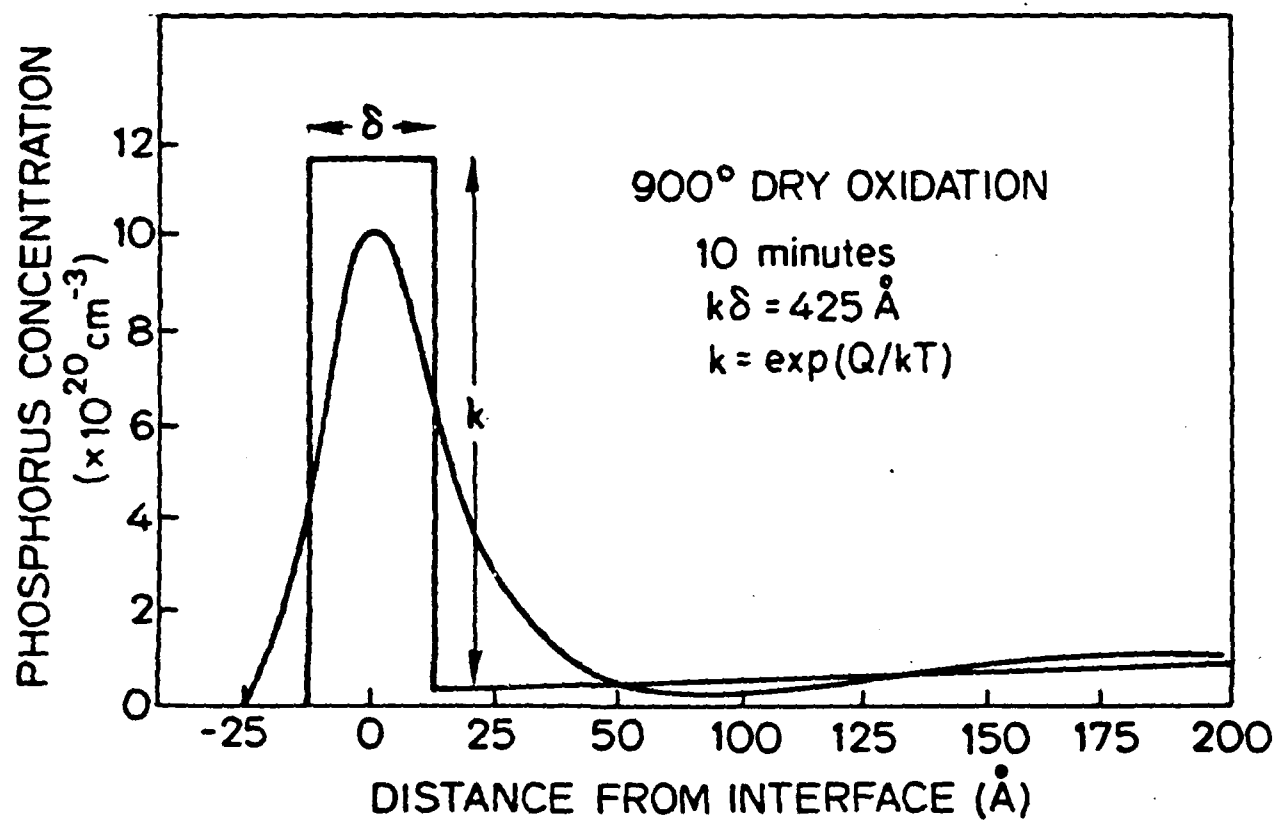
Fig. 16.10: A comparison between our initial transient theory and an Auger sputter profile obtained from a sample oxidized for ten minutes at 900°C. In the theoretical curve, δ has been arbitrarily set at 25 A. The parameter k · δ was determined from measurements on equilibrium distributions at 900°C.

415
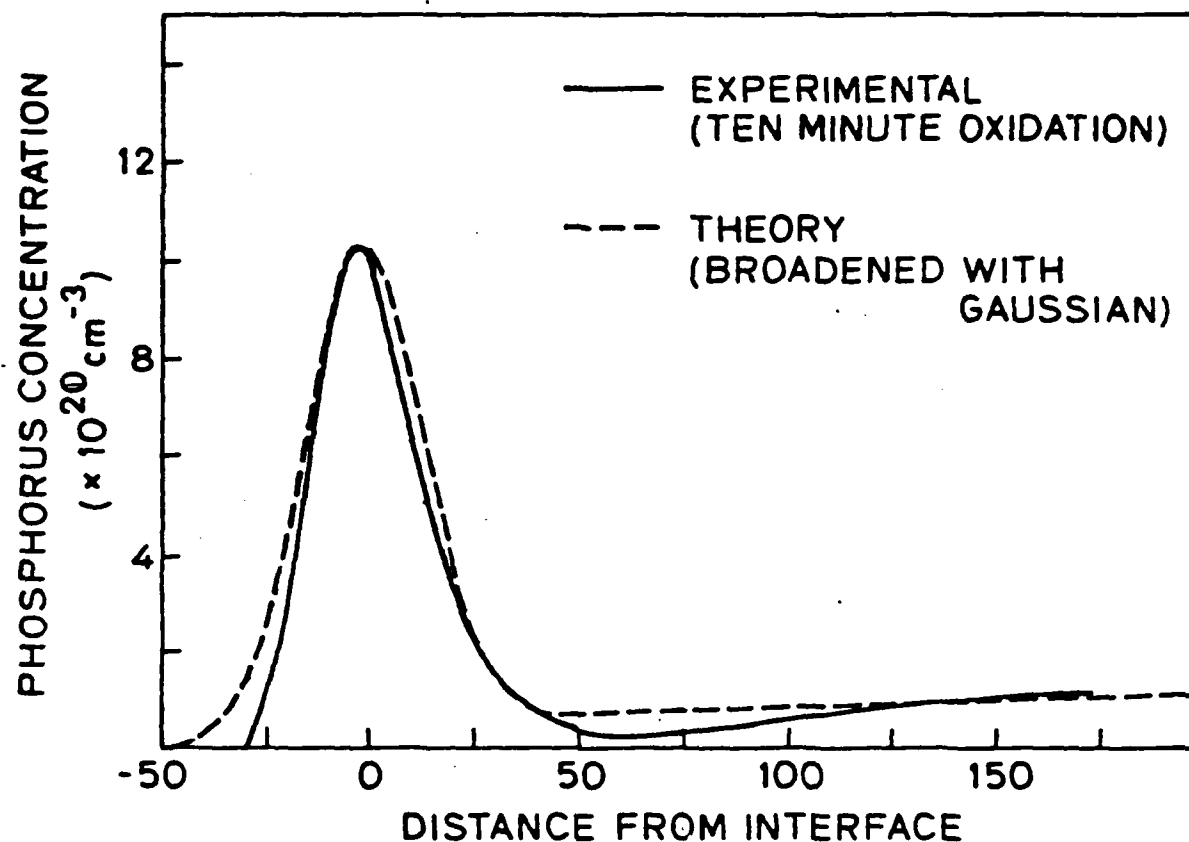
Fig. 16.11: Comparison betweeen an Auger sputter profile, and a theoretical curve that has been convoluted to account for ion knock-on mixing. The theoretical curve was obtained with $k \cdot \delta = 425$ Å, $\delta = 3$ Å, and then convoluted with a gaussian distribution (standard deviation = 14 Å).
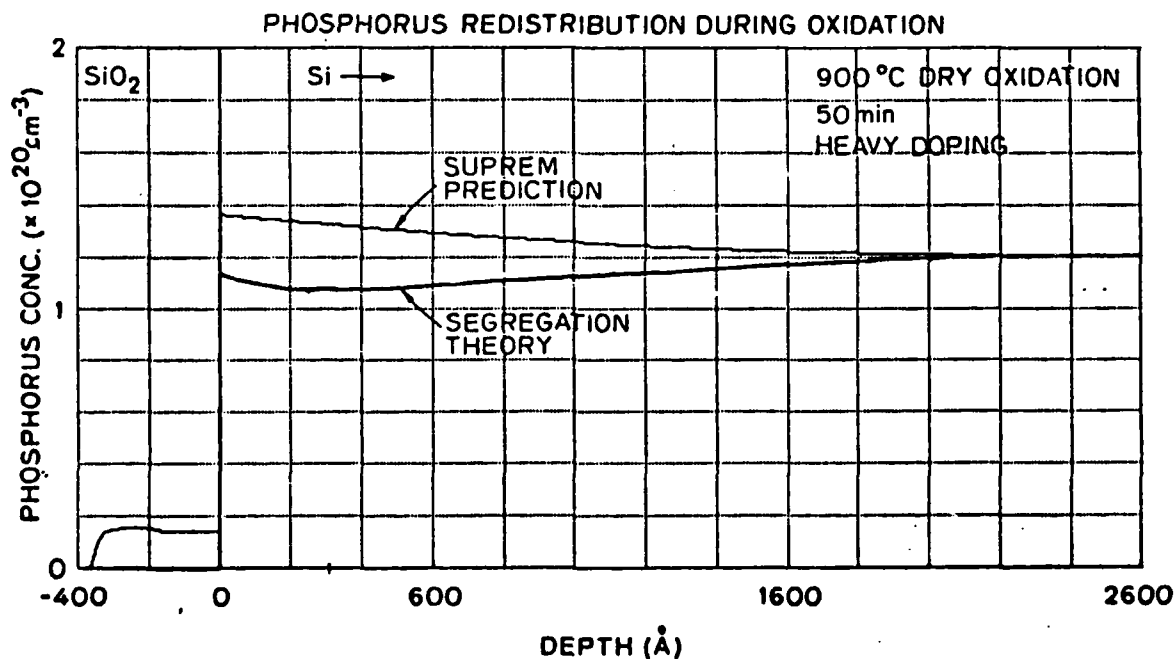
PHOSPHORUS REDISTRIBUTION DURING OXIDATION

Fig. 16.12: Comparison between our initial transient theory (which accounts for interface segregation) and that redicted from the SUPREM II computer program (which does not account for segregation). Both predictions assume an initial uniform doping of 10^{20} cm^{-3}.
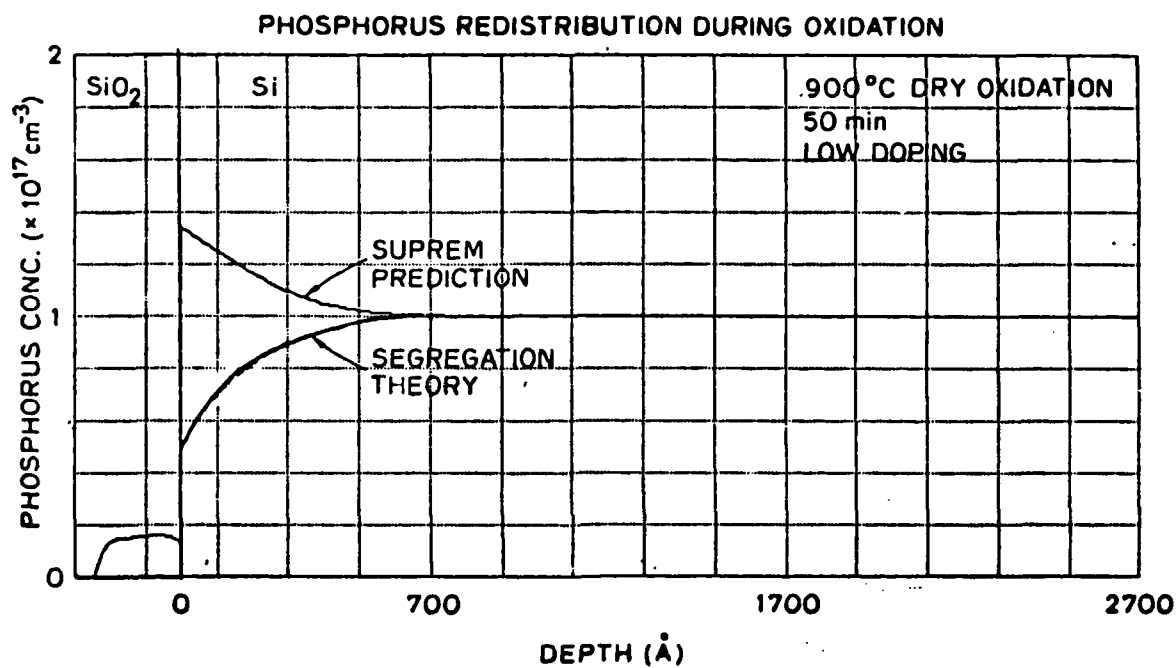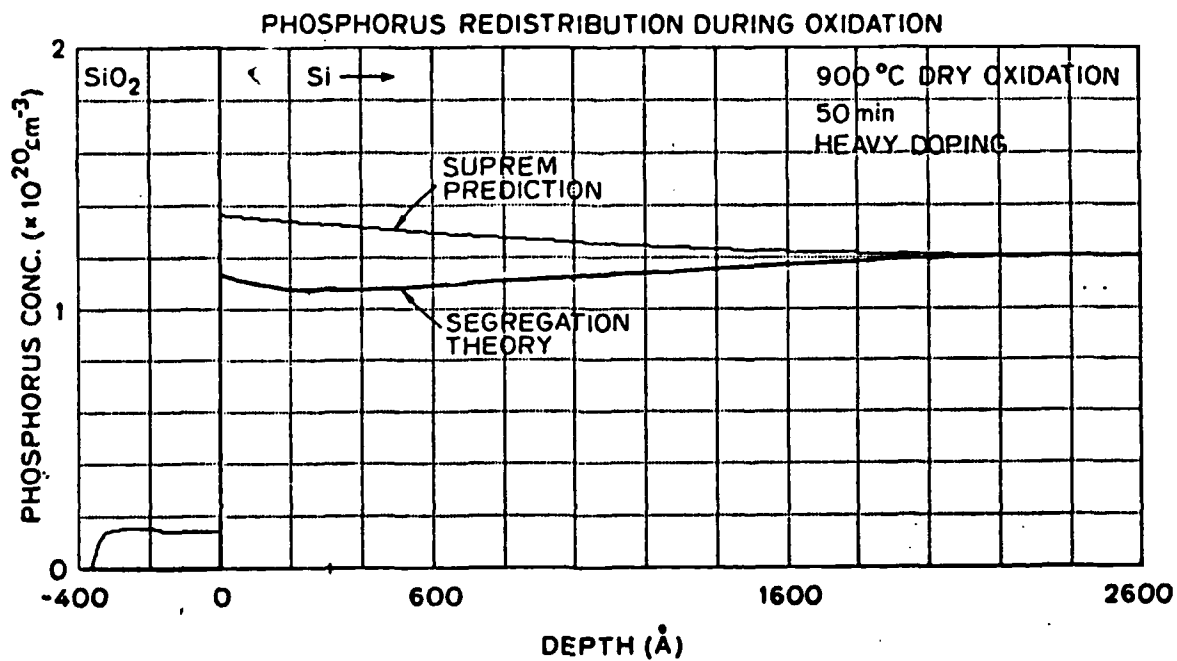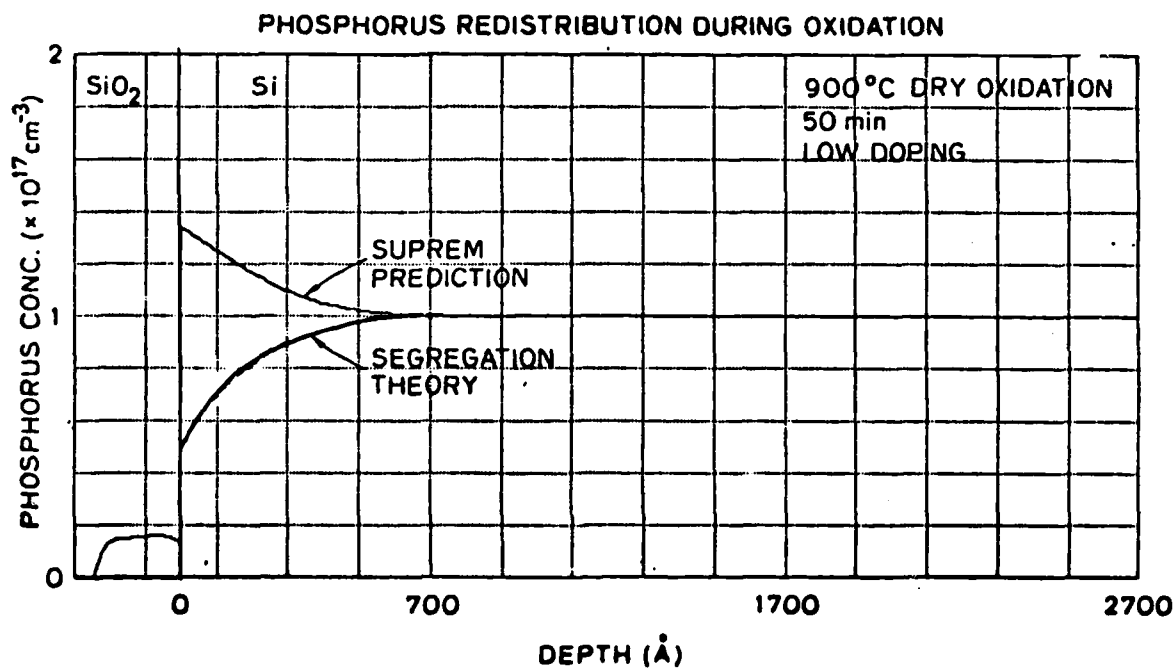
417

PHOSPHORUS REDISTRIBUTION DURING OXIDATION

Fig. 16.13: Comparison between our initial transient theory and that predicted from the SUPREM II computer program. Both predictions assume an initial uniform doping of 10^{17} cm^{-3}

PHOSPHORUS REDISTRIBUTION DURING OXIDATION

Fig. 16.12: Comparison between our initial transient theory (which accounts for interface segregation) and that redicted from the SUPREM II computer program (which does not account for segregation). Both predictions assume an initial uniform doping of 10^{20} cm^{-3}.

PHOSPHORUS REDISTRIBUTION DURING OXIDATION

Fig. 16.13: Comparison between our initial transient theory and that predicted from the SUPREM II computer program. Both predictions assume an initial uniform doping of 10^{17} cm^{-3}
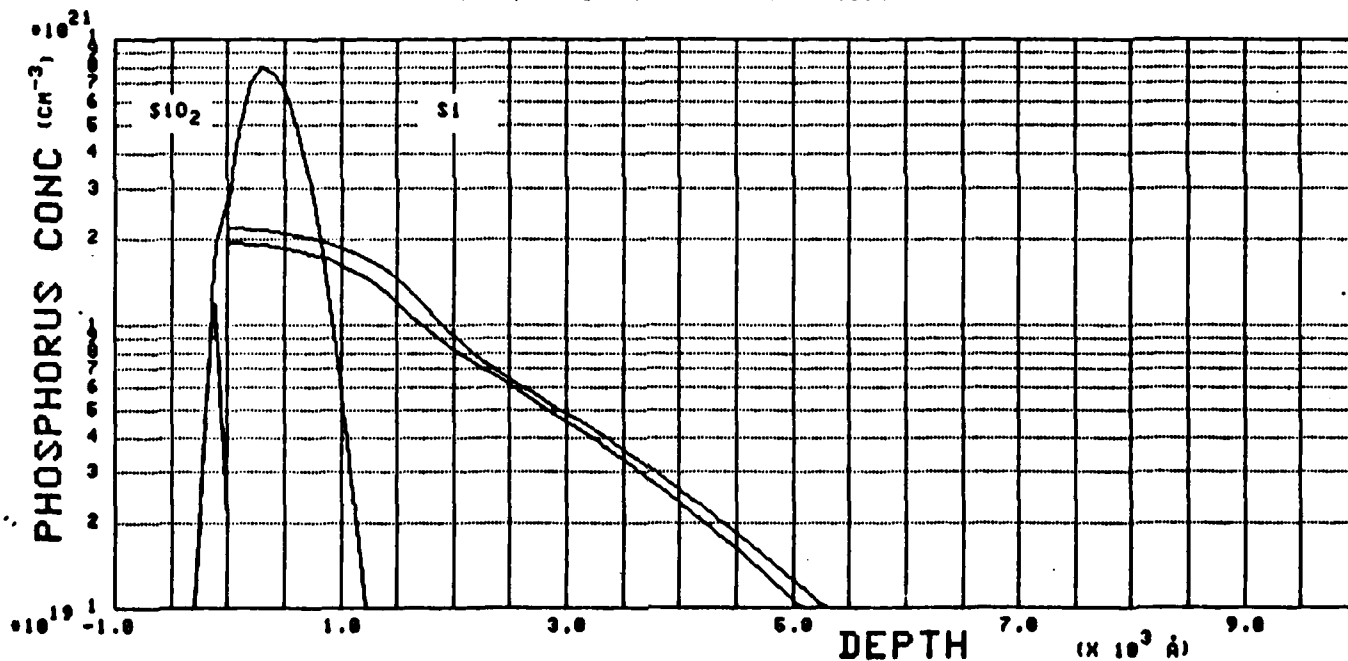
PHOSPHORUS IMPLANT AND DIFFUSION

Fig. 16.14: Approximate effect of interface segregation on diffused profiles from an ion implanted layer. Both profiles have been annealed in N_2 for 120 minutes at 900°C. One profile is lower in concentration, however, because 6×10^{14} cm^{-2} of its dopant has segregated to the Si/SiO$_2$ interface. The implanted dose was 5×10^{15} cm^{-2}.

17.   EFFECTS OF ION BEAM MIXING ON PROFILES OF THE Si-SiO$_2$ INTERFACE

FOR THIN OXIDES

M. A. Taubenblatt and C. R. Helms

17.1   INTRODUCTION

Previous Auger sputter profiling studies of ultrathin SiO$_2$ layers [1] have shown that thinner layers result in the reduction of the Si-SiO$_2$ interface widths. These results have been explained by intrinsic properties of the thin oxide layers or by surface roughening effects [17.1,17.2,17.3]. In this study, a model is developed, based on the spatial distribution of the ion energy, which can be modeled by a position dependent diffusion constant. This sputter induced diffusion causes the observed interface broadening, and is greatly reduced for thin oxides, due to the kinetics of the process in accordance with experimental observations. To verify this model, in addition to the data of references 1-3, additional samples were investigated to insure that the growth process was not effecting the interface width. This was done starting with a 100 Å oxide that was chemically etched thinner so that the interface in all cases would be identical.

In section II, this model is presented, and an approximate analytical expression is developed which gives the dependence of interface width on the initial oxide thickness. Also exact numerical solutions of the diffusion equation are calculated, which give the shape of the Auger profile, and show that the analytical solution is correct.

In section III, experimental procedure is discussed and the techniques used to measure the initial oxide thickness. These were; ellipsometry, sputtering time to the interface (50% point) and the peak heights of the Auger KLL transitions.

In section IV the results are compared to theory.

17.2 THEORETICAL MODEL

The derivation of the relationship between the deposited ion energy and the effective diffusion coefficient, D, follows the analysis of Schwartz and Helms and also Haff and Switkowski [17.4,17.5]. We repeat this derivation here for the sake of clarity.

The model is analogous to thermal diffusion theory. A bombardment ion penetrates the surface, causing random collisions, and thus displacements of the bulk atoms just beyond the surface. For low ion energies the scattering is isotropic, thus the assumption of random displacements is valid. The path of an atom is described by a random walk. The net distance R traveled by this atom is then

$$R^2 = n\bar{a}^2 t \tag{17.1}$$

where $\bar{a}^2$ is the mean square displacement length and n is the number of displacements occuring per atom per unit time. The diffusion constant is then given by [17.6]

$$\frac{\frac{1}{6} \partial R^2(t)}{\partial t} = \frac{1}{6} n\bar{a}^2 \tag{17.2}$$

To determine n, the Kinchin-Pease formula is used [17.7]. This formula gives the number of collisions occuring due to an ion of energy E as $E/2E_d$ where E_d is a displacement energy. The number of collisions is thus E/E_d since two atoms collide in a displacement event.

At this point we depart from previous analyses, and consider the ion to deposit its energy over a spatial range. Thus let $E(Z)$ be the energy deposited per unit depth into the target. For low ion energies, the displaced atoms due to one ion will not stray far from their initial locations, and one can then consider $E(Z)/E_d$ to be the local number of collisions per unit length caused by an incident ion at depth Z into the target. Thus $n(Z)$ is given by:

$$n(Z) = E(Z) \, J/E_d en \qquad (17.3)$$

where J is the incident ion current, e the electron charge, and n the atom density of the target. It then follows from (17.2) and (17.3) that

$$D(Z) = \frac{1}{6} \frac{E(Z)J\bar{a}^2}{E_d en} \qquad (17.4)$$

For low ion energies $\bar{a}^2$ is roughly given by the nearest neighbor distance squared, $(3A)^2$, and E_d for SiO_2 is 25 eV. It only remains then to determine the energy deposition profile $E(Z)$. This profile can be approximated using the LSS theory [17.8] based on using scattering (stopping) cross sections to determine the moments of energy and ion distributions. The moments for a power law (1/3) scattering, suitable for low energy ions is calculated in Sigmund and Sanders [17.9] and is used here to estimate the moments $\langle x \rangle$ and $\langle \Delta X^2 \rangle$ of the energy distribution, where x is the direction of ion incidence. Then $E(x)$ takes the functional form

$$E(x) = \frac{e^{-(x-\langle x \rangle)^2}}{2 \langle \Delta x^2 \rangle} \qquad (17.5)$$

which must be normalized, so that

$$E_{ion} = \int_0^\infty E(x)dx \qquad (17.6)$$

where E_{ion} is the incident ion energy.

The mechanism of interface width reduction with thin oxides is now clear. Figure 1 shows $E(x)$ plotted for a thick and a thin oxide. For a thick oxide, the entire energy distribution will sweep past the interface region as the target is sputtered away. But for a thin oxide, much of the energy is deposited beyond the interface, so that the total energy deposited near the interface is reduced.

We may now define an approximate broadening parameter

$$\bar{W}^2 = (Dt)_{total} \tag{17.7}$$

If D were a constant, $\bar{W}^2$ would be an exact solution of the broadening, but since D is dependent on both position and time (due to sputtering) (17.7) is only approximately correct. Thus we may write

$$\bar{W}^2 = \int_0^d Ddt = \int_0^d D\frac{dx}{r} = \int_0^d \frac{1}{6} \frac{E(x)\bar{a}^2}{E_d S} dx \tag{17.8}$$

where d = oxide thickness, r = sputtering rate = $\frac{JS}{en}$ and S = sputtering yield in atoms/ion.

When d is large compared to the extent of E(x) (d >> <x>) $\bar{W}^2$ reduces to

$$\bar{W}^2_{max} = \frac{E_{ion}a^2}{6 \ E_d S} \tag{17.9}$$

which is identical with [17.4]. Using (17.5), and evaluating (17.8) in terms of error functions, we arrive at the equation;

$$\bar{W}^2(d) = \bar{W}^2_{max} \left[\frac{\mathrm{erf}\ \alpha + \mathrm{erf}\left(\frac{d}{\sqrt{2}\ <\Delta x^2>} - \alpha\right)}{1 + \mathrm{erf}\ \alpha} \right] \tag{17.10}$$

where $\alpha = <x>/\sqrt{2<\Delta x^2>}$. This equation is plotted in Fig. 17.2 with $\alpha = 1.3$ (α is nearly constant over a variety of ions and is independent of energy).

In order to verify that the approximation (17.7) is correct, the numerical solution of the diffusion equation with position (and time) dependent D has been calculated using finite difference equations. Fig. 17.3 compares the numerical solution to an error function solution with Dt given by (17.10) for an initial step function profile. The numerical solution is very nearly an error function, and the approximation (17.7) is thus shown to be quite accurate.

The above discussion applied to ions normally incident to the target. In our experimental chamber the ions strike the target at 49° off normal incidence. Since the 2nd moment in the transverse direction ($<y^2>$) in non-zero, this causes an additional complication. Fig. 17.3 shows an off-normal incident ion beam.

The x direction is along the incident ion trajectory, the y direction transverse to this, and the z direction along the target normal. To compute E(Z) (energy deposition as a function of depth into the target), we must integrate along the y direction. Thus E(Z) is given by

$$E(Z) = J \int_{-\infty}^{z/\sin\theta} \exp\left[-(x-\langle x\rangle)^2/2\langle\Delta x^2\rangle - y^2/2\langle y^2\rangle\right] dy \qquad (17.11)$$

where $x = z/\cos\theta + y\tan\theta$ and K = normalization constant. Substituting for x, and defining $\bar{z} = \langle x\rangle \cos\theta$, we have

$$E(Z) = K \int_{-\infty}^{z/\sin\theta} \exp\left[-(\frac{z - \bar{z}}{\cos\theta} - y\tan\theta)^2/2\langle\Delta x^2\rangle - y^2/2\langle y^2\rangle\right] \qquad (17.12)$$

This may be reduced to the equation

$$E(Z) = Ke^{-(x - \bar{x})^2/2 \langle\Delta x^2_{eff}\rangle} \frac{\sqrt{\pi}}{2A}\left[1 + erf\left(\sqrt{A}\left(\frac{z}{\sin\theta} - \frac{B}{A}(z-\bar{z})\right)\right)\right] (17.13)$$

where $A = (\tan^2\theta/\langle\Delta x^2\rangle + 1/\langle y^2\rangle)$

$B = \tan\theta/\cos\theta \langle\Delta x^2\rangle$

$1/\langle\Delta z^2 eff\rangle = (B^2/A + 1/\cos\theta \langle\Delta x^2\rangle)$ \qquad (17.14)

For typical parameters the term in brackets varies $\approx$10% as z goes from 0 to ∞, so that it can be neglected. Thus, for off normal incidence, E(z) has approximately the same functional form as for normal incidence, but with a new $\langle\Delta z^2 eff\rangle$ as defined in (17.14).

To calculate the moments needed to determine E(z) we use the computations of Sigmund and Sanders [17.4], representing the SiO_2 lattice as a solid with atomic mass 20 (average mass of SiO_2) with the atomic density of SiO_2. A 1/3 power law potential is used, which is valid over the range 100-7000 eV for Ne and 300-22,000 eV for Ar. These parameters are listed in Table I.

17.3 EXPERIMENTAL

The oxide samples were prepared on (100) n-type starting material, and grown in ~1% O_2, 99% Ar at 1000°C for ~90', using standard integrated circuit techniques. The oxide grown in this manner was 100 A. The thinner oxide samples were prepared by etching this 100 A oxide in 200:1 HF solution (etch rate ~15A/min). Thus any effects due to oxide growth kinetics were eliminated.

The oxide thicknesses prepared in the above manner were measured using three techniques; ellipsometry, sputtering time to 50% point, in the profiling experiment and the ratio of the Auger KLL lines for Si (1606 eV) and SiO_2 (1619 eV) for the starting surface.

The ellipsometry was performed using a Gaertner automated ellipsometer and Gaertner software. The ellipsometric measurements of the oxides were quite sensitive to the assumed substrate index of refraction (especially the imaginary part). The generally used value of 3.85 - 0.02i did not give results in agreement with other techniques. Since the substrate index is known to vary with surface preparation [17.10], we found much better agreement by measuring the substrate index for our samples directly. Less than 15" elapsed between removing the sample from 100:1 HF and the measurement. The substrate indices measured in this way ranged from Ns = 3.83 to 3.85 and Ks = -0.08 to 0.16.

Thicknesses from sputter times were determined by using a pre-determined sputter rate from thicker oxides measured with ellipsometry. Taking sputtering geometry and escape length effect into account, the sputtering time to the 50% point of the Si 92 eV Auger line is 2.5 A less than the original interface, assuming a near step interface.

Due to escape length effects, the KLL Auger line of Si and SiO_2 (1609 eV, 1619 eV) can be used to determine the thickness of oxides less than ~50A. The Auger signal strengths for Si and SiO_2 are given by:

$$S_{Si} = e^{-d/L}$$

$$S_{SiO_2} = A(1 - e^{-d/L}) \tag{15}$$

where d = oxide thickness, L = escape length at 1600 eV = 22 and A = the senstivity ratio of bulk SiO_2 to Si = 2. Figure 5 compares the thicknesses determined by the above techniques. Auger sputter profiling (ASP) was done in a Varian ASP system using a single pass CMS (cylindrical mirror analyzer) with integrated electron gun, and a Varian ion gun mounted at 49% off sample normal.

17.4 COMPARISON WITH THEORY

To compare this theory of interface broadening with experiment, we have measured the interface width (10% to 90%) for oxide samples of various thicknesses for 3KeV Ne and 1 KeV Ar.

For an error function profile the 10% to 90% width is given by $W_{knock-on} = 3.6 \; \bar{W}^2$, with $\bar{W}^2$ defined by (7) and (10). We must also consider the broadening effect of escape length, given by 2.2L [17.11] and the intrinsic interface width, W_0. These broadening factors add in an RMS manner [17.11], so that

$$W_{measured} = \sqrt{W_0^2 + W_{knock-on}^2 + (2.2L)^2} \tag{17.16}$$

Thus we compute a $W_{knock-on}$ experimental $= \sqrt{W_{measured}^2 - W_0^2 - (2.2L)^2}$ (17.17)

The comparison of theory and experiment is shown in Fig. 6 where Wo is assumed to be 10 . The assumptotic thick oxide, value of $W_{knock-on}$ has been compared in [17.4], and since we are interested in the behavior of thin oxides compared to the generic function (17.10), the data curves have been fitted to W_{max}.

As can be seen, the data follows the predicted curves, with 3 KeV Ne having a much more extended effect than 1 KeV Ar, since the energy deposition curve for 3 KeV Ne penetrates more quickly.

17.5 CONCLUSION

A theory has been developed, relating the interface broadening to the energy deposition profile of the sputtering ions analgous to thermal diffusion. From this theory, we expect thinner oxides to have narrower interfaces, and for this effect to be more pronounced for deeply penetrating ions, such as 3 KeV Ne. Compared to 1 KeV Ar, analysis of oxide samples of identical starting thickness shows good agreement with this theory. This data is not explained well by assuming thin oxides have narrower intrinsic interfaces, since all starting oxides were of identical thickness, nor by roughening theories, since they do not account for the differing behavior of 3 KeV Ne and 1 KeV Ar.

TABLE I

Ion	Energy	$\langle x \rangle$	$\langle x^2 \rangle$	$\langle y^2 \rangle$	$\bar{z}$	$\langle \Delta z_{eff}^2 \rangle$
Ne	500 eV	16Å	$(10Å)^2$	$(8Å)^2$	10Å	$(5Å)^2$
Ne	1000 eV	25Å	$(16Å)^2$	$(12Å)^2$	18Å	$(9Å)^2$
Ne	3000 eV	52Å	$(33Å)^2$	$(24Å)^2$	34Å	$(18Å)^2$
Ar	500 eV	11Å	$(7Å)^2$	$(5Å)^2$	7Å	$(4Å)^2$
Ar	1000 eV	18Å	$(11Å)^2$	$(8Å)^2$	12Å	$(6Å)^2$
Ar	3000 eV	37Å	$(24Å)^2$	$(17Å)^2$	24Å	$(13Å)^2$

REFERENCES

[17.1] J. F. Wager and C. W. Wilmson, *J. Appl. Phys.*, 50, 874 (1979).

[17.2] C. F. Cook, C. R. Helms, D. C. Fox, *J. Vac. Sci. Technol.*, 17, 44, (1980).

[17.3] H. Frenzel, P. Balk. The Physics of MOS Insulators Proceedings of the International Topical Conference, Raleigh, North Carolina, June 18-20, 1980.

[17.4] S. A. Schwartz and C. R. Helms, *J. Vac. Sci. Technol.*, 16, 781, (1979).

[17.5] P. K. Haff, Z. E. Switkowski, *J. Appl. Phys.*, 48, 3383 (1977).

[17.6] F. Reif, Statistical Thermodynamics, McGraw-Hill, 448, (1965).

[17.7] G. H. Kinchkin, R. S. Pease, *Rep. Progr. Phys.*, 18, 2, (1955).

[17.8] J. Lindhard, M. Scharff, H. E. Schiott, *Mat. Fys. Medd. Dan. Vid. Selsk*, 33, 14 (1963).

[17.9] P. Sigmund, J. B. Sanders, Int. Conf. on Appl. of Ion Beams to Sem. Tech. Grenoble, 214, (1967).

[17.10] K. Vedam and Samuel S. So, *Surf. Sci.*, 29, 379 (1972).

[17.11] C. R. Helms, N. M. Johnson, S. A. Schwarz, W. E. Spicer, *J. Appl. Phys.*, 50, 7007 (1979).
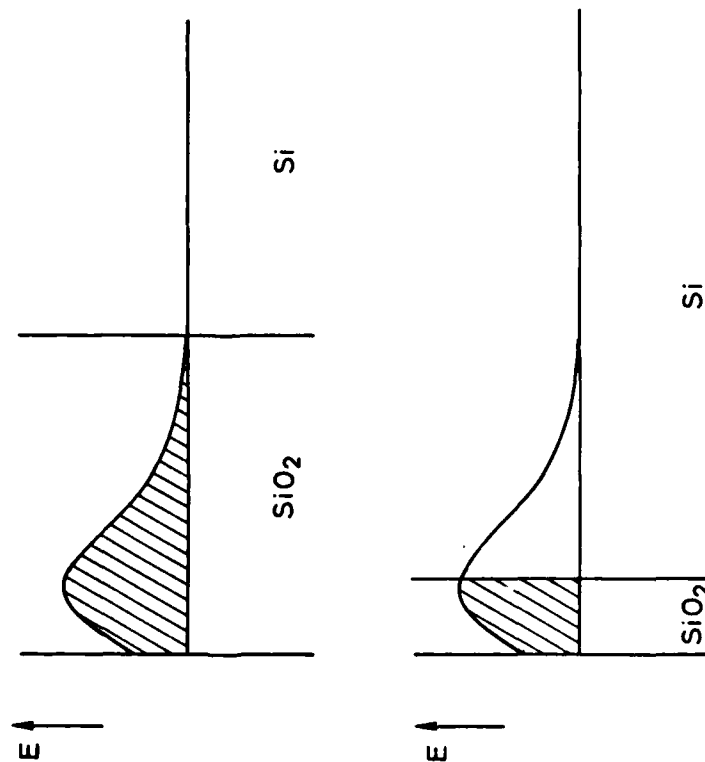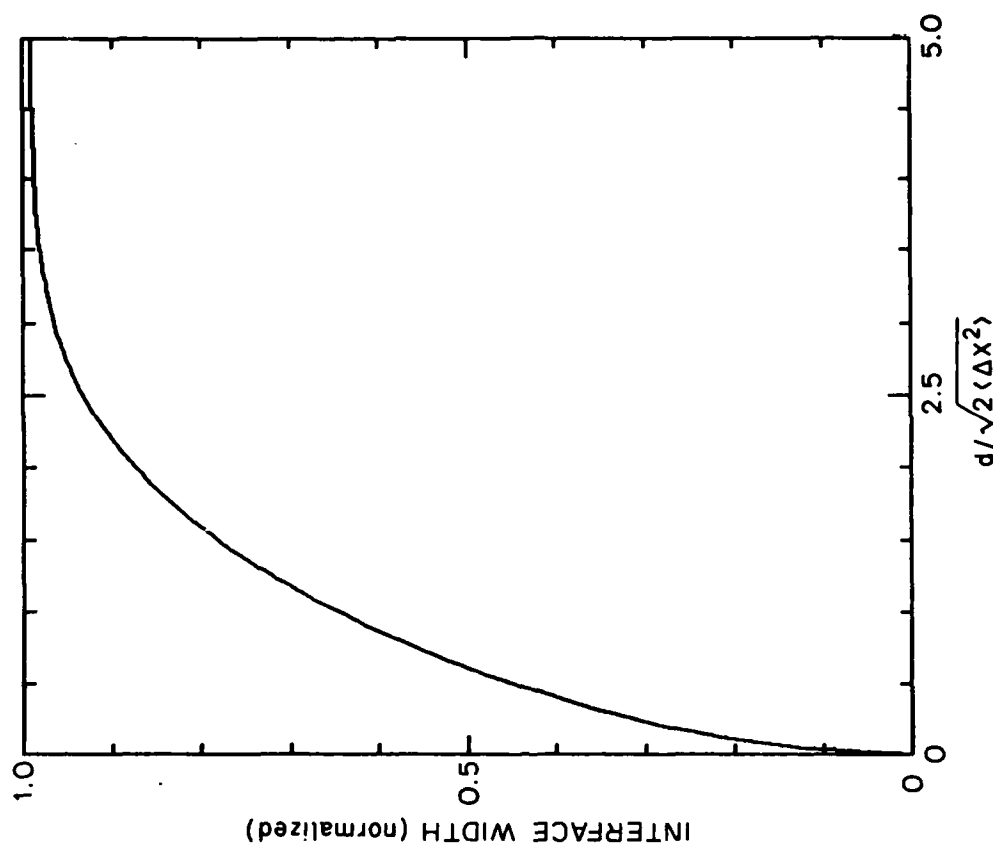
Fig. 17.2: The general function for interface width, $\sqrt{W^{-2}/W_{max}^{-2}}$ versus $d\sqrt{2\langle\Delta x^2\rangle}$, where d = initial oxide thickness.

Fig. 17.1: Schematic diagram of the spatial distribution of the ion energy for thick and thin SiO_2 layers on Si.
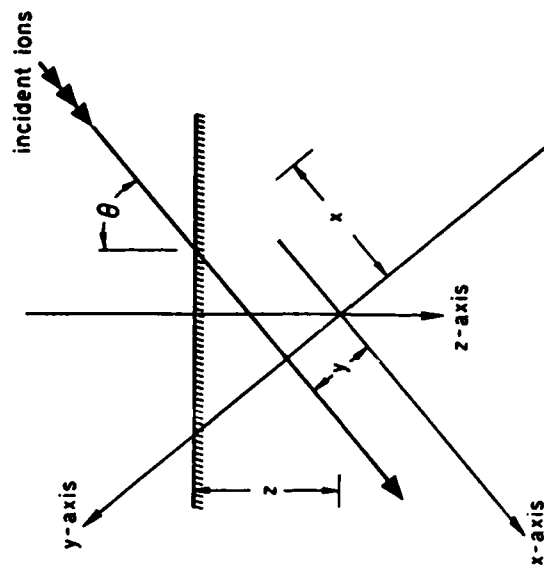
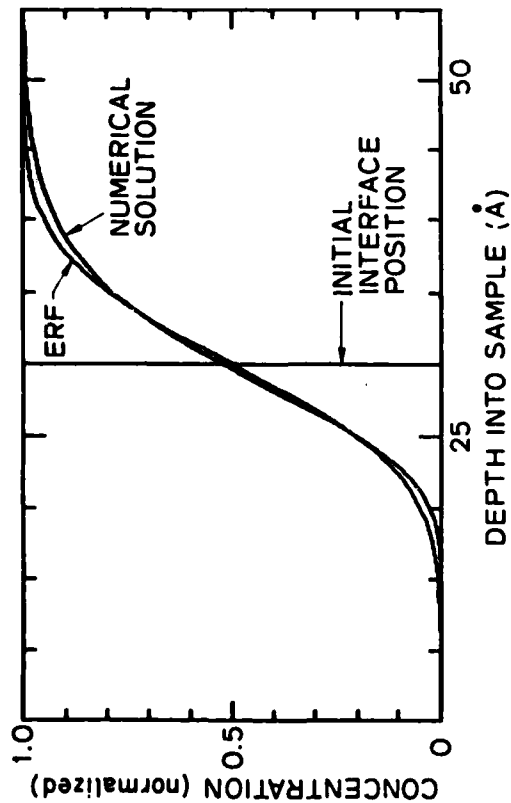Fig. 17.4: Diagram of axes for computation of parameters for an off-normal ion beam.

Fig. 17.3: Numerical solution to the diffusion equation compared to an error function.
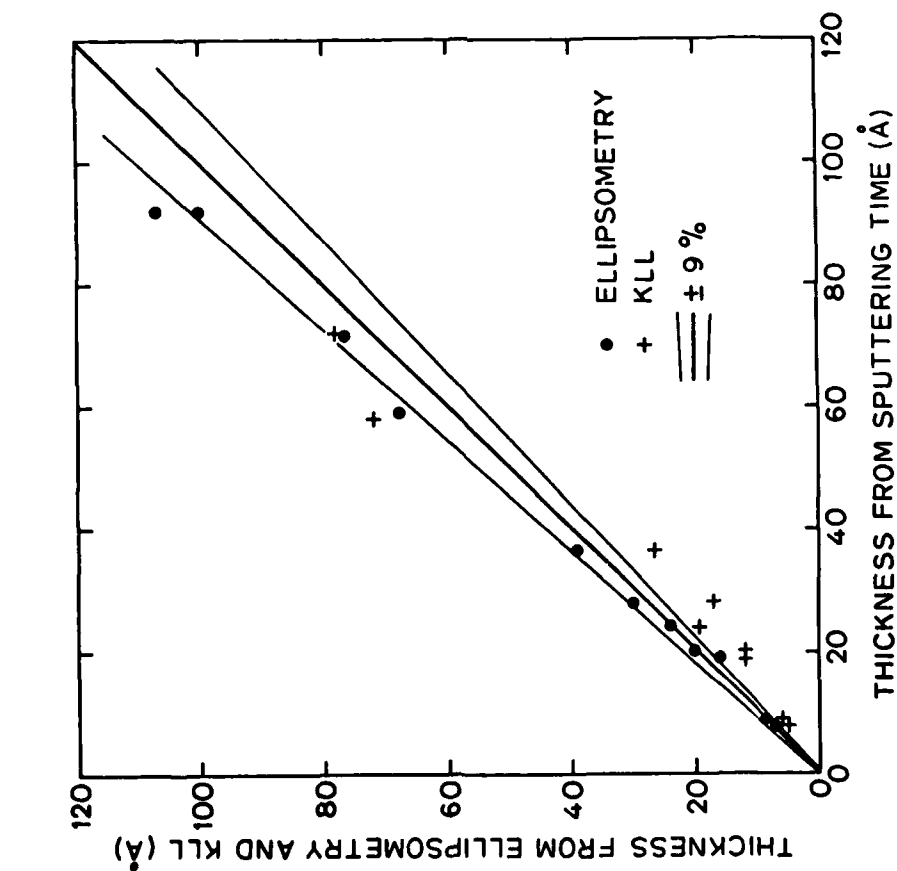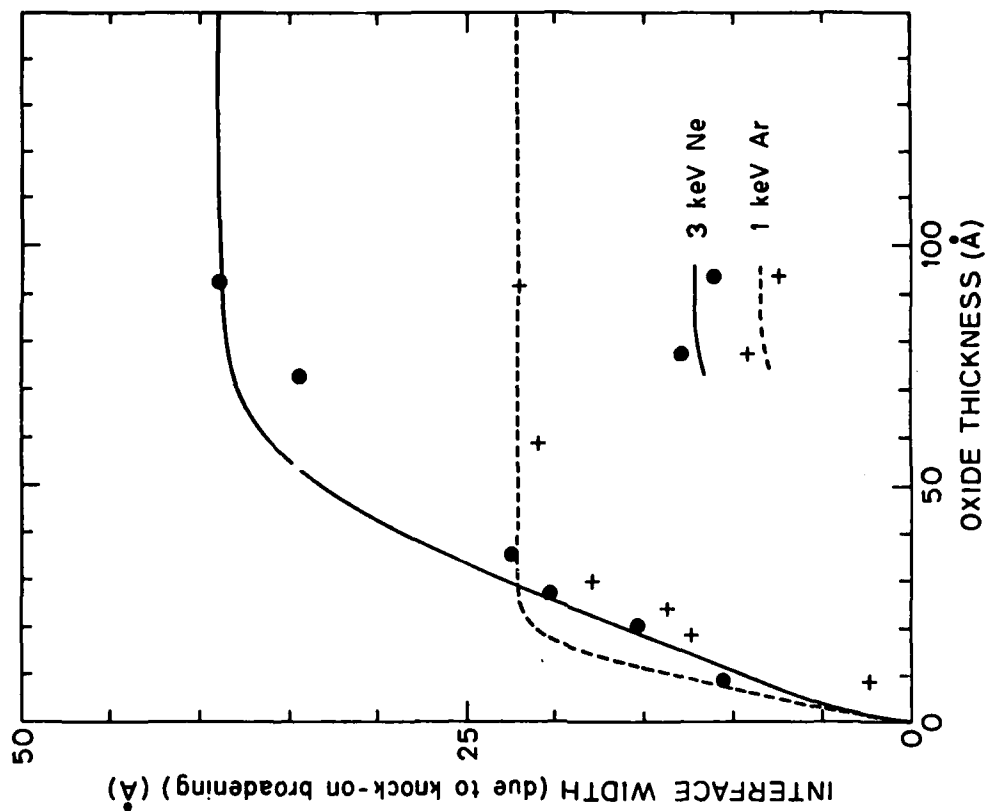
Fig. 17.6: Comparison of theory versus experiment,
Error in interface width and oxide
thickness is ±9%. Oxide thickness is
determined from the sputter time.

Fig. 17.5: Comparison of thicknesses determined
by ellipsometry, sputtering time,
and KLL Auger line ratios.

18.   Si3N4 OXIDATION KINETICS

T. I. Kamins

18.1 INTRODUCTION

It has long been recognized that a model for the oxidation of silicon nitride would be a useful addition to SUPREM since silicon nitride is often oxidized during the formation of a field oxide using LOCOS processing (Fig. 18.1). It is necessary to know the oxidation rate both to ensure that a sufficient thickness of nitride is available to mask the field oxidation and also to enable etching of the grown oxide over the nitride without excessive thinning of the field oxide outside the nitride-covered regions.

In the past year it was decided to address the subject of oxidation of nitride, but it was recognized that manpower limitations would not allow a complete experimental study. Instead it was decided to attempt to compile existing data from the literature and from SUPREM users; if sufficient data could be obtained, a preliminary model would be included in SUPREM III, with refinement of the model as further information became available. Initially it was desired to consider the oxidation of silicon nitride deposited both by atmospheric-pressure and low-pressure chemical vapor deposition in the 900-950°C temperature range normally used for these formation processes. Silicon nitride formed by plasma-enhanced CVD at lower temperatures was not to be considered in this initial study. Information about the oxidation at both atmospheric pressure and at high pressures was to be obtained, if possible, since high-pressure oxidation is becoming more commonly used for the formation of thick field oxides.

The study started with a thorough literature search, which produced information about the oxidation of silicon nitride under a wide range of
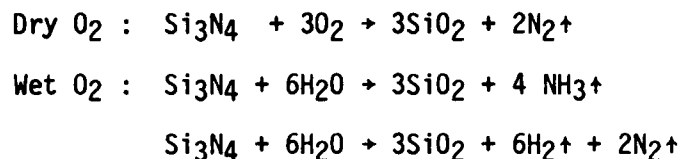
432

experimental conditions. In each experiment the nitride was formed under
somewhat different conditions and processed differently so that direct compar-
ison between studies is somewhat questionable. Since, however, any model used
in SUPREM will initially be applied to material formed and processed under
varied conditions, it was decided to ignore the different experimental condi-
tions in the various studies and attempt to obtain a fairly consistent model
for as much of the experimental data as possible. Some of the experimental
data was rather arbitrarily excluded if it differed markedly from the majority
of the available data or if the authors, when they could be contacted, ex-
pressed severe reservations about the quality of the data. Data from refer-
ences [18.1-18.7] was used in the study, while data fom references [18.8-
18.11] was excluded.

In addition to obtaining information from the open literature, a survey
was sent to SUPREM users requesting any data which they might have on the oxi-
dation of silicon nitride. It was felt that this application would be a
useful test of the concept of compiling data for construction of a model since
the fine details of the experimental conditions were not to be considered.
The data obtained from the survey were very limited. One set of data was vol-
untarily supplied by Gordon Wright of Synertek. Another (and very valuable)
set of data was obtained from Jim Stimmell of National. This turned out to be
some of the most useful information in the entire study since it contained
data taken under consistent conditions at several different pressures and
allowed modification of the tentative model.

18.2 POSSIBLE OXIDATION MECHANISMS

The basic possible oxidation mechanisms are shown in Table 18.1 below for
dry or wet oxidation. ("Wet oxidation," "steam oxidation," and "pyrogenic oxi-
dation" will be lumped together in this discussion.) In both cases a gaseous

TABLE 18.1 - POSSIBLE REACTIONS

Dry O_2 : $Si_3N_4 + 3O_2 \rightarrow 3SiO_2 + 2N_2\uparrow$

Wet O_2 : $Si_3N_4 + 6H_2O \rightarrow 3SiO_2 + 4 NH_3\uparrow$

$Si_3N_4 + 6H_2O \rightarrow 3SiO_2 + 6H_2\uparrow + 2N_2\uparrow$

byproduct is formed along with the solid silicon dioxide. It is the presence
of this byproduct and the necessity of removing it from the silicon nitride-
silicon dioxide interface which can cause possible differences from the well-
characterized oxidation of silicon. Figure 18.2a shows the fluxes of the
oxidizing species and the nitrogen-containing byproduct, while the concentra-
tions of these species are shown in Fig. 18.2b. In our analysis we will
consider that the transfer of the oxidizing species at the outer surface
of the oxide layer is rapid so that the concentration of the oxidizing species
in the oxide near its outer surface is proportional to the pressure of the
water vapor in the gas phase, and the concentration of the nitrogen containing
byproduct near the outer surface of the oxide is zero.

If there is no effect of the nitrogen-containing species at the
oxide-nitride interface, the analysis is analogous to the oxidation of
silicon. Expressions for the diffusion and reaction fluxes can be written:

$$F_I = D_0 \frac{C_{os} - C_{oi}}{X_0} ; \quad F_2 = 6 k_n C_{oi} \qquad (18.1)$$

In steady state, the fluxes may be equated and related to the formation
rate of the oxide:

$$2N \left(\frac{d_{X_0}}{dt}\right) = F = \frac{6 k_n C_{os}}{1 + 6 k_n X_o/D_o} \qquad (18.2)$$

434

This equation may be solved for the oxide thickness:

$$X_0^2 + \left(\frac{D_0}{3 k_n}\right) X_0 = \left(\frac{D_0}{N}\right) C_{os} \, t \qquad (18.3)$$

which is in the familiar form

$$x_0^2 + A x_0 = B t \qquad (18.4)$$

where the linear rate constant B/A in the reaction-rate-limited case may be written as

$$\frac{B}{A} = \frac{3 k_n}{N} C_{os} = \frac{3 k_n H}{N} P \qquad (18.5)$$

and the parabolic rate constant B in the diffusion-limited case is

$$B = \frac{D_0}{N} C_{os} = \frac{D_0 H}{N} P \qquad (18.6)$$

Both rate constants are proportional to the pressure P of the water vapor in the gas phase. A log-log plot of oxide thickness vs. time should show a slope of unity at short times and a slope of 1/2 for long times. Alternatively the product of pressure and time may be used as the independent variable.

If however, the oxidation rate is affected by the concentration of the nitrogen-containing species in the system, the analysis becomes more complicated. If the diffusion of the nitrogen species away from the nitride-oxide interface is slow, its presence there may impede the reaction rate at this interface. In addition, the presence of this species in the oxide may retard the inward diffusion of the oxidizing species.

We can consider the first of these effects by following a model by Enomoto, et al [18.2]. As before, we can write expressions for the fluxes of the

435

oxidizing species related to diffusion and surface reaction and equate these expressions in steady state:

$$F_o = D_o \frac{C_{os} - C_{oi}}{X_o} = 6k_n C_{oi} \qquad (18.7)$$

In addition, we must now write an expression for the similar fluxes related to the nitrogen-containing byproduct:

$$F_n = D_n \frac{C_{ni} - C_{ns}}{X_o} = 4k_n C_{oi} \qquad (18.8)$$

The crucial assumption in Enomoto's model is made at this point. He assumes that the reaction-rate "constant" is, indeed, constant when the concentration of the nitrogen-containing bypoduct at the interface is small so that it does not impede the surface reaction. Above a critical concentration, the reaction rate constant is assumed to decrease inversely with increasing concentration of the byproduct at the interface. This assumption is admittedly arbitrary but allows analytical solution of the equations in a special case of interest. With this assumption, the oxidation rate can be written as

$$\frac{DX_o}{dt} = \frac{9 \, k_n' \, C_n' \, D_n'}{4 \, D_o \, N} \left[\sqrt{1 + \frac{4 \, D_o^2 \, C_{os}}{9 \, k_n' \, C_n' \, D_n \, X_o}} \; -1 \right] \qquad (18.9)$$

in the region where the surface-reaction rate is impeded by the presence of the byproduct at the interface. In the surface-reaction-rate-limited region of oxidation, the second term in the square root is dominant, and the oxide thickness can be obtained:

$$X_o = \left[\frac{81 \, k_n' \, D_n \, C_n' \, C_{os}}{16 \, N^2} \right]^{1/3} t^{2/3} = \left[\frac{81 \, k_n' \, D_n \, C_n' \, H}{16 \, N^2} \right]^{1/3} p^{1/3} \, t^{2/3} \qquad (18.10)$$

The oxide thickness is found to be proportional to the 2/3 power of the time. For longer times, the second term is small, and the solution reduces to the

familiar diffusion-limited case with a 1/2 power dependence on time. For small times, of course, the linear rate "constant" is assumed to be constant so that the linear behavior seen in the simpler case is observed. Therefore, on a log thickness-vs-log time plot (Fig. 18.3), a region with slope 2/3 should appear between the region with slopes of unity and 1/2. Whether this region can be differentiated from the gradual transition from a slope of 1 to a slope of 1/2 in the simpler model is debatable, although Enomoto argues that this region can be unambiguously recognized. It should also be noted that Enomoto's model predicts a 1/3 power dependence of oxide thickness on the pressure of the water vapor.

In any study of the oxidation of silicon nitride, the relationship between the oxide thickness formed and the amount of nitride consumed must be known. The widely varying values given in the literature arise because of the different values used for the density of silicon nitride. Values between 2.8 and 3.4 g/cm^3 have been used. The most consistent observations of the dependence of the density on the deposition conditions are reported by Chu, et al. [18.12]. They report an increase in density from 2.78 to 3.01 g/cm^3 as the deposition temperature is increased from 850 to 1200°C, compared to a bulk density for silcon nitride of 3.18 g/cm^3. If the deposition or final annealing temperature is the dominant variable determining the density, a value of approximately 2.8 g/cm^3 is applicable in the cases of most practical interest, with a corresponding conversion ratio of 1.6 - 1.7.

18.3 EMPIRICAL MODEL

With these preliminary comments, we can analyze the available data to obtain a model useful for SUPREM. A tentative examination of the data indicated that, over a limited range, a simple relationship could be used to fit the available data. As an initial approximation we can assume an inverse

exponential dependence on temperature and power-law dependences on time and water-vapor pressure:

$$\Delta X_n = K_1 \; e^{-Ea/kT} \; t^{K_2} \; p^{K_3} \tag{18.11}$$

If Enomoto's model is valid over this range, the powers of time and pressure should be 2/3 and 1/3, respectively. If the simpler linear-parabolic model is applicable, the product of time and pressure could be used as the independent variable, with the power varying between 1 and 1/2. Figure 18.4 shows the dependence of nitride consumed on the product of time and pressure at 1000 °C, where the most extensive data is available. If the nitride consumed is dependent only on the product, the data should fall on a single line. Since it obviously does not, the different model must be used.

From the data of Fig. 18.4, however, we do see that the amount of nitride consumed depends approximately on the 0.7 power of time so that K2 = 0.7 in the empirical model. As we will see below, the 0.7-power dependence is also reasonable for temperatures of 900 and 1100°C. In order to obtain the pressure dependence, the time required to consume a constant amount of nitride can be found as a function of the pressure. A plot of log time vs. log pressure should be a straight line with a slope of -K3/K2. Fig. 18.5 shows that this relationship holds over the pressure range 1-10 atm. The slope of -1.42 indicates that K3 = 0.7 x 1.42 = 1.0; that is, the amount of nitride consumed should be linearly proportional to the pressure, unlike the 1/3-power dependence predicted by Enomoto's model. The ratio of the amount of nitride consumed to the pressure is plotted as a function of time in Fig. 18.6 for a temperature of 900°C, indicating the usefulness of the relationship over this range of time, pressure, and temperature although the physical mechanism responsible for the linear dependence is not clear. Similar plots are obtained

438

at 1000 and 1100°C. Finally, the apparent activation energy can be found from an Arrenhius plot using data from Fig. 18.6 and similar data at other temperatures at fixed times. From Fig. 18.7, we can see that an activation energy of approximately 1.9 eV satisfies the data. There is, however, more uncertainty in this parameter than in the others since the choice of data to emphasize is somewhat arbitrary. If data only from the study by Miyoshi, et. al., which covers the widest temperature range, is used, a value of 1.74 eV is obtained. This suggests that extrapolation to widely different temperatures is not valid. In addition, a plot of very limited data from Miyoshi, et. al., at 800°C suggests that different mechanisms may be dominant at lower temperatues, since the time-dependence corresponds to a slope closer to unity than to the value of 0.7 applicable at higher temperatures.

Within the empirical model, however, the values of Ea, K1, K2, and K3 can be specified. IF the value of activation energy of 1.9 eV is used, K1 is found to be 1.1×10^7, while K1 is 2.3×10^7 if the activation energy is chosen to be 1.74 eV. The final empirical expressions can then be written as

$$\Delta X_n = 1.1 \times 10^7 \left(e^{-1.9 \text{ eV}/kT} \right) t^{0.7} \, P_{H_2O}$$

or (18.12)

$$\Delta X_n = 2.3 \times 10^7 \left(e^{-1.74 \text{ eV}/kT} \right) t^{0.7} \, P_{H_2O}$$

where the amount of nitride consumed is given in nanometers, time is in minutes, and water-vapor pressure is in atmospheres. Comparison with experimental data shows reasonable agreement over the temperature range 900 to 1100°C and the pressure range 1 to 10 atm. The model does not seem to be accurate at 800°C; data taken at much higher pressures (eg 70 atm [18.10]) or at pressures much less than one atmosphere [18.2] are also not in agreement.

More data is needed, especially at lower temperatures, which will be used in practical high-pressure oxidation in the near future.

Within these limitations, however, the empirical model should be useful as a preliminary model in SUPREM III and has been implemented. It can be refined and extended, hopefully from user feedback. In addition, the time, pressure, and temperature dependences identified in this study may be useful in identifying the physical mechanisms controlling the oxidation of silicon nitride.

REFERENCES

[18.1]   D. R. Craven and J. B. Stimmell, Semiconductor International (June 1981), 59.

[18.2]   T. Enomoto, R. Ando, H. Morita, and H. Nakayama, Japan J. Appl. Phys., 17, 1049 (1978).

[18.3]   H. Miyoshi, N. Tsubouchi, and A. Nishimoto, J. Electrochem. Soc., 125, 1824 (1978).

[18.4]   I. Franz and W. Langheinrich, Solid-State Electron., 14, 499, (1971).

[18.5]   A. P. Turley, R. M. McLouski, P. R. Reid, and D. H. McCann, J. Electrochem. Soc., 123, 117, (1976).

[18.6]   Y. Tamaki, S. Isomae, S. Mizuo and H. Higuchi, J. Electrochem. Soc., 128, 644 (1981).

[18.7]   E. L. MacKenna, private communication.

[18.8]   J. A. Appels, E. Kooi, M. M. Paffen, J. J. H. Schatorje, and W. H. C. G. Verkuylen, Philips Res. Repts., 25, 118 (1970).

[18.9]   A. K. Gaind and E. W. Hearn, J. Electrochem. Soc., 125, 139, 1978).

[18.10] R. J. Powell, J. R. Ligenza, and M. S. Schneider, IEEE Trans. Electron. Dev., ED-21, 636 (1974).

[18.11] S. I. Raider, R. Flitsch, J. A. Aboaf, and W. A. Pliskin, J. Electrochem. Soc., 123, 560, (1976).

[18.12] T. L. Chu, C. H. Lee, and G. A. Gruber, J. Electrochem. Soc., 114, 717 (1967).
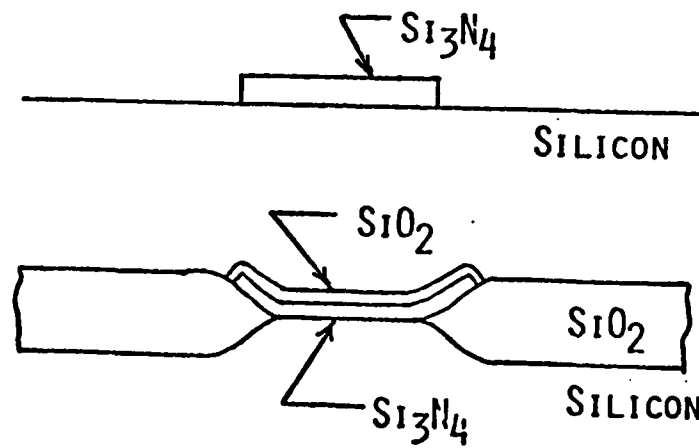
LOCOS STRUCTURE

Fig.18.1: Illustration of LOCOS (local oxidation) process for which Si_3N_4 is commonly used.
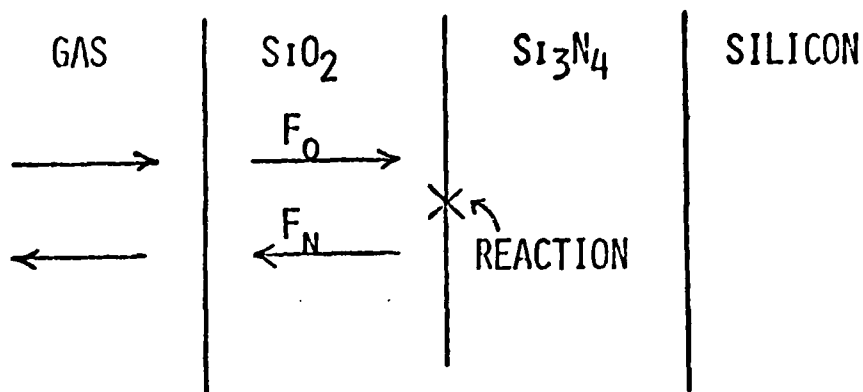
Fig.18.2a: Fluxes of oxidizing species and the nitrogen-containing by-product in Si_3N_4 oxidation.
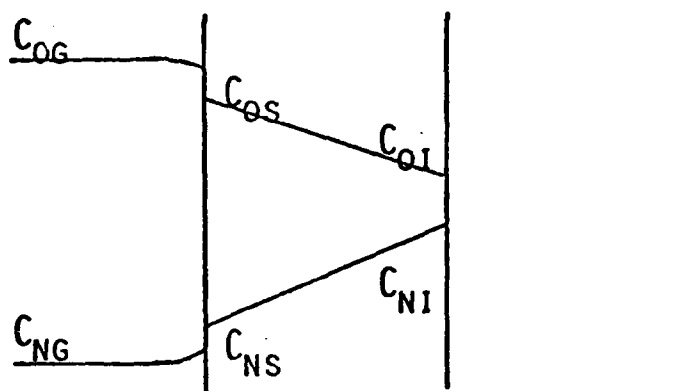

Fig. 18.2b: Concentration profiles of the reactants and reaction products illustrated in Fig. 18.2a.
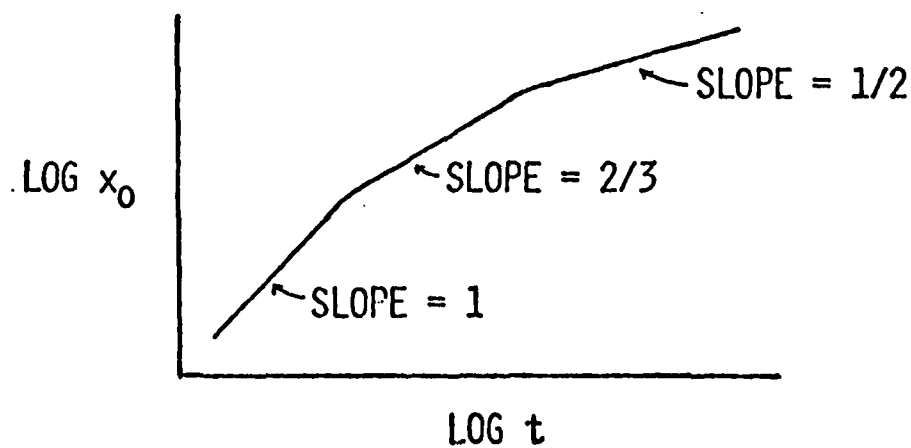

Fig.18.3: Qualitative predictions of Enomoto's model for the growth kinetics of SiO_2 on Si_3N_4.
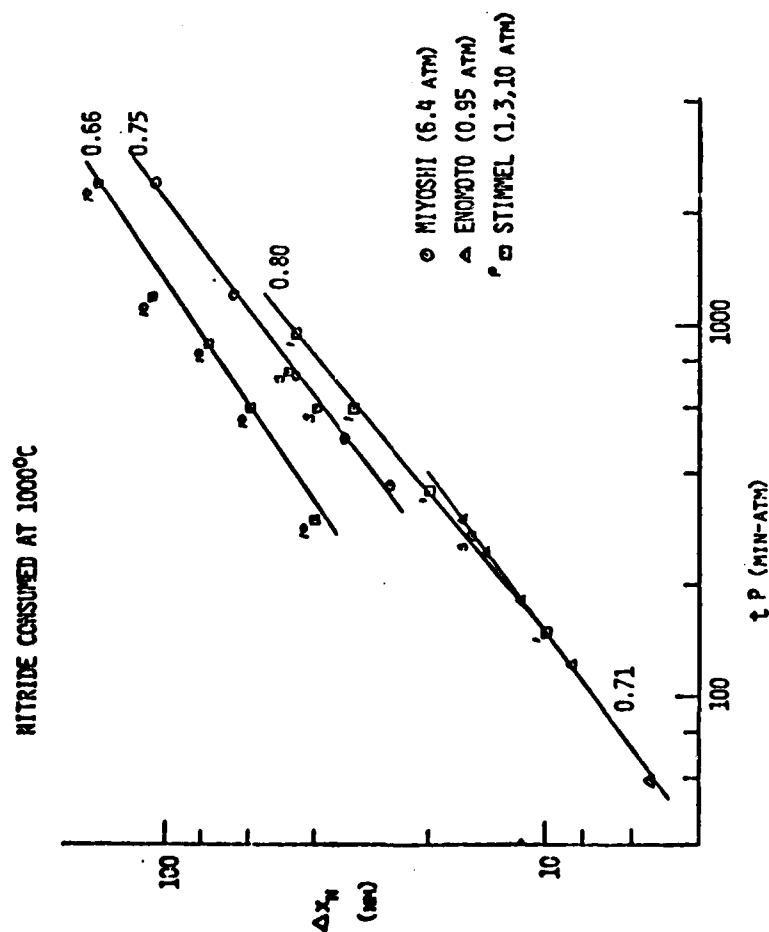
TIME FOR $\Delta x_N = 40$ NM vs PRESSURE

SLOPE = -1.42

DATA: STIMMEL

PRESSURE (ATM)

t (MIN)

Fig.18.5: Time required to consume 40 nM of Si_3N_4 vs. H_2O oxidant pressure at 1000°C.

NITRIDE CONSUMED AT 1000°C

0.66
0.75
0.80

⊙ MIYOSHI (6.4 ATM)
▲ ENOMOTO (0.95 ATM)
▪ STIMMEL (1,3,10 ATM)

0.71

Δx_N (NM)

t P (MIN-ATM)

Fig.18.4: Dependence of nitride consumed on the product of time and pressure of oxidation at 1000°C.
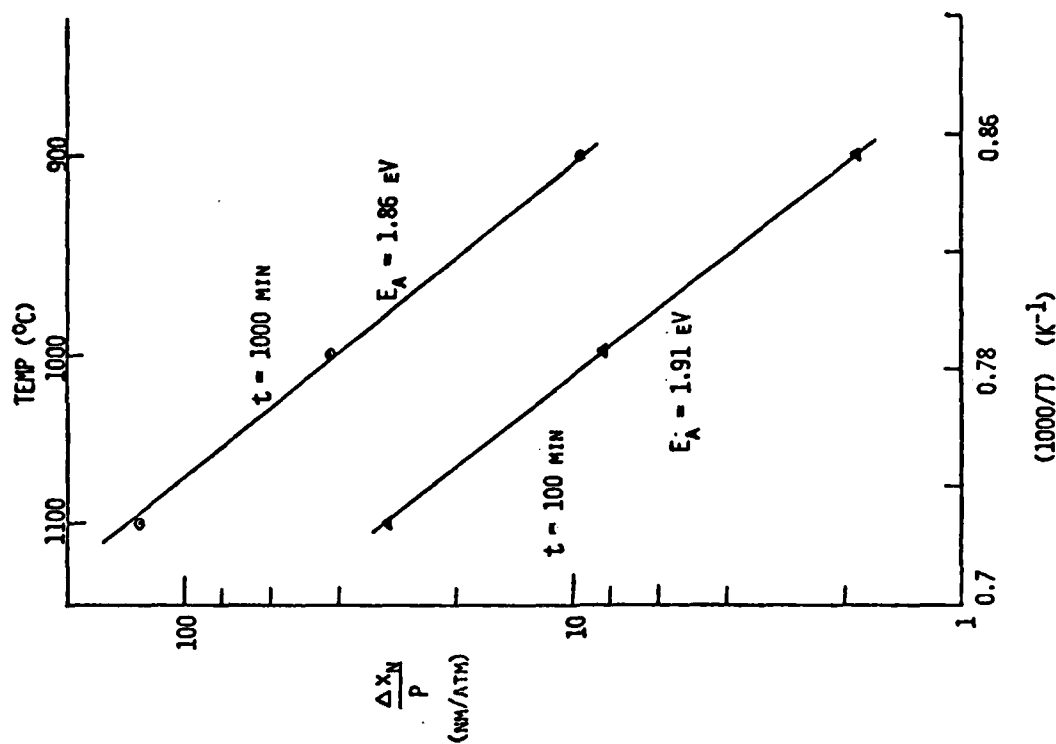
Fig.18.7: Ratio of amount of Si_3N_4 consumed to the oxidant pressure vs. $1/T$, which yields the activation energy.
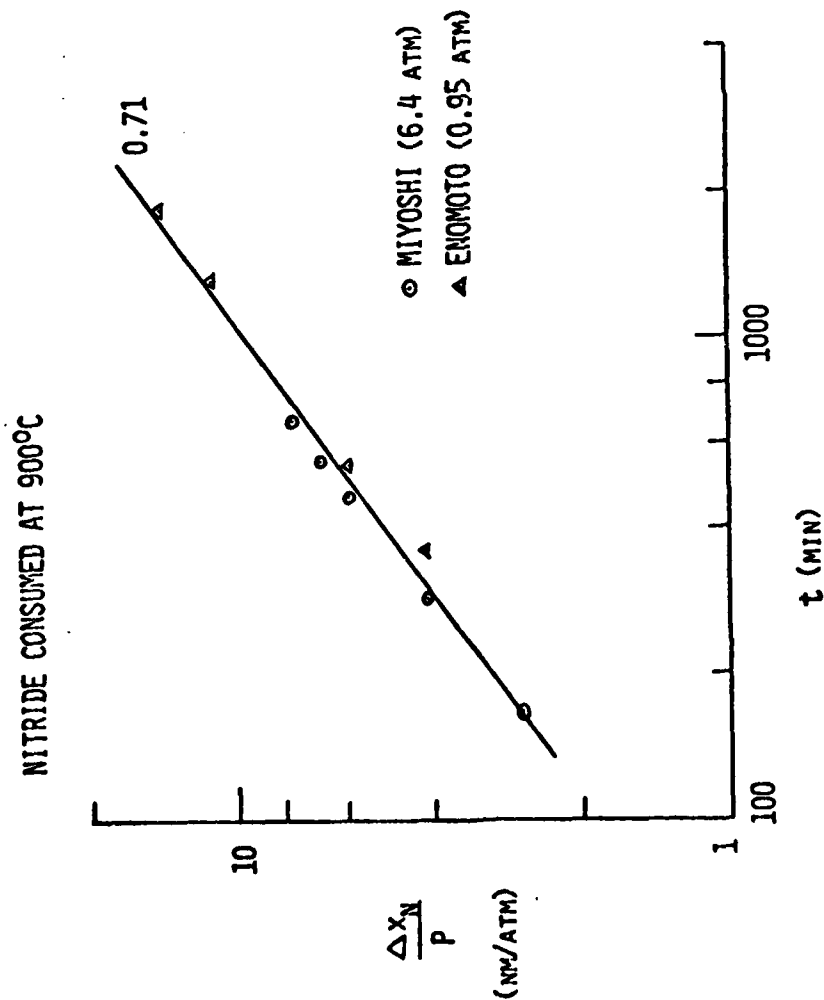
Fig.18.6: Ratio of amount of Si_3N_4 consumed to the oxidant pressure, vs. time at 900°C.

444